

UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE CIÊNCIAS ECONÔMICAS
CEDEPLAR – Programa de Pós-Graduação em Demografia

Igor Coura de Mendonça

**DIFERENCIAIS DEMOGRÁFICOS EM ÍNDICES MULTIDIMENSIONAIS DE
VULNERABILIDADE: estudo de caso do impacto do desastre em Mariana nas famílias de
trabalhadores em condição de pobreza**

Belo Horizonte

2022

Igor Coura de Mendonça

**DIFERENCIAIS DEMOGRÁFICOS EM ÍNDICES MULTIDIMENSIONAIS DE
VULNERABILIDADE: estudo de caso do impacto do desastre em Mariana nas famílias de
trabalhadores em condição de pobreza**

Dissertação apresentada ao Programa de Pós-Graduação em Demografia do Centro de Desenvolvimento e Planejamento Regional da Faculdade de Ciências Econômicas da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do Título de Mestre em Demografia.

Orientador: Professor Doutor Gilvan Ramalho Guedes

Coorientadora: Professora Doutora Ana Maria Hermeto Camilo de Oliveira

Belo Horizonte

2022

Ficha catalográfica

M539d
2022

Mendonça, Igor Coura de .
Diferenciais demográficos em índices multidimensionais de vulnerabilidade [manuscrito]: estudo de caso do impacto do desastre em Mariana nas famílias de trabalhadores em condição de pobreza / Igor Coura de Mendonça. – 2022.
117f.: il., tabs.

Orientador: Gilvan Ramalho Guedes.
Coorientadora: Ana Maria Hermeto Camilo de Oliveira.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Centro de Desenvolvimento e Planejamento Regional.
Inclui bibliografia.

1. Companhias de mineração – Teses. 2. Vítimas de desastres – Teses. 3. Impacto ambiental – Teses I. Guedes, Gilvan Ramalho. II. Oliveira, Ana Maria Hermeto Camilo de. III. Universidade Federal de Minas Gerais. Centro de Desenvolvimento e Planejamento Regional. IV. Título.

CDD: 304.6

Elaborada por Rosilene Santos CRB6-2527
Biblioteca da FACE/UFMG. RSS – 004/2023



ATA DE DEFESA DE DISSERTAÇÃO DE IGOR COURA DE MENDONÇA - NÚMERO DE REGISTRO 2019650767.

Às treze horas do dia dezoito do mês de outubro de dois mil e vinte e dois, reuniu-se na Faculdade de Ciências Econômicas da Universidade Federal de Minas Gerais a Comissão Examinadora de DISSERTAÇÃO, indicada *ad referendum* pelo Colegiado do Curso em 26/09/2022, para julgar, em exame final, o trabalho final intitulado “Diferenciais demográficos em índices multidimensionais de vulnerabilidade: estudo de caso do impacto do desastre em Mariana nas famílias de trabalhadores em condição de pobreza”, requisito final para a obtenção do Grau de Mestre em Demografia, área de concentração em Demografia. Abrindo a sessão, o Presidente da Comissão, Prof. Gilvan Ramalho Guedes, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra ao candidato, para apresentação de seu trabalho. Seguiu-se a arguição pelos examinadores, com a respectiva defesa do candidato. Logo após, a Comissão composta pelos professores Gilvan Ramalho Guedes, Ana Maria Hermeto Camilo de Oliveira, Raquel Zanatta Coutinho e Mariangela Furlan Antigo se reuniu, sem a presença do candidato e do público, para julgamento e expedição do resultado final. A Comissão aprova o candidato por unanimidade. O resultado final foi comunicado publicamente ao candidato pelo Presidente da Comissão. Nada mais havendo a tratar o Presidente encerrou a reunião e lavrou a presente ATA, que será assinada por todos os membros participantes da Comissão Examinadora. Belo Horizonte, 18 de outubro de 2022.

Prof. Gilvan Ramalho Guedes
(Orientador) (CEDEPLAR/FACE/UFMG)

Profa. Ana Maria Hermeto Camilo de Oliveira
(Coorientadora) (CEDEPLAR/FACE/UFMG)

Profa. Raquel Zanatta Coutinho
(CEDEPLAR/FACE/UFMG)

Profa. Mariangela Furlan Antigo
(CEDEPLAR/FACE/UFMG)

PROFA. LAURA LÚCIA RODRÍGUEZ WONG
Coordenadora do Programa de Pós-Graduação em Demografia

Dedico este trabalho ao Gabriel, minha alegria diária; à Alice, minha flor de cerejeira; e À Karol, que cumpriu sua promessa de dez anos atrás.

AGRADECIMENTOS

À Bernadete, ao Paulo, à Isabel, ao Bruno, à Lorena e à Dione, por me acolherem na hora mais escura.

Aos professores Barbieri, Wong, Cássio, Simone, Irineu e Bernardo, por me orientarem para além deste trabalho.

Ao BDMG, especialmente à Marta, à Polyana, à Áurea, ao Jorge, ao Amabile, pela paciência, pelo suporte e por desejarem o melhor para mim.

À Professora Hermeto, pela orientação e pela inspiração, que vem desde muito antes de entrar no Cedeplar.

Ao Professor Gilvan, pelo conhecimento, pelo respeito, pela insistência e pela empatia.

Ao Kender, por ter me perguntado diariamente o que eu iria estudar naquele dia.

À Alice e ao Gabriel, por terem acabado com a minha paz para me trazer alegria.

À Karol, por ter ficado comigo na alegria e na tristeza, na saúde e na doença, na riqueza e na pobreza, amando-me.

Obrigado a todos pelo privilégio de ter tanta fonte de resiliência ao meu redor.

Sofri muito nesse período, mas o meu sofrimento foi para prosperar. Só posso imaginar como é sofrer para resistir. Faço votos de que os anônimos que geraram as minhas estatísticas sejam felizes e prósperos.

“(…) enquanto houver lugares onde seja possível a asfixia social; em outras palavras, e de um ponto de vista mais amplo ainda, enquanto sobre a terra houver ignorância e miséria, livros como este não serão inúteis.” (Victor Hugo, Os miseráveis, 1862)

RESUMO

O rompimento da barragem de Fundão em Mariana (MG) demonstra como a demografia pode ser fator crucial nas capacidades das famílias de se recuperarem ou de agravarem suas condições. Partindo de uma base de dados desidentificada do Cadastro Único do Ministério da Cidadania (CadÚnico), esta dissertação teve como objetivo principal mostrar o potencial que indicadores multidimensionais de vulnerabilidade social possuem em capturar os efeitos de um desastre tecnológico nas famílias atingidas, inclusive identificando diferenciais agravantes ou mitigadores na composição familiar ante mudanças súbitas e adversas nas condições de acesso ao trabalho e de escassez de recursos ao longo dos anos após o desastre. A partir do estudo de caso do desastre de Mariana, foi elaborada uma avaliação de impacto por diferenças em diferenças, tomando como referência seus efeitos no Índice de Vulnerabilidade das Famílias (IVF): um indicador multidimensional de vulnerabilidade social. Procurou-se comparar os resultados anuais no indicador global e nas dimensões de trabalho e de renda, primeiro pela população total do CadÚnico, e em seguida analisando apenas o universo de famílias com certas condições demográficas: presença de filho nascido vivo; presença de idoso; ausência de cônjuge; e minoria do domicílio em idade ativa. O grupo de tratamento engloba os 35 municípios mineiros oficialmente considerados atingidos no desastre. Já para elencar os municípios do grupo de controle foi empregado o algoritmo *K-Nearest Neighbor*, assumindo 23 variáveis como parâmetro para parear aqueles mais semelhantes aos 35 atingidos, obtendo como resultado um total de 169 classificados. A avaliação se dividiu em períodos, considerando cada um deles como um número de anos a partir de dezembro de 2014 (onze meses antes do rompimento), tendo como curto prazo o período de 2014-15; médios prazos 2014-16 e 2014-17; e longo prazo 2014-2018. Este trabalho apresentou resultados aparentemente ambíguos, mas que conseguem traçar um cenário interessante para se analisar e debater. Considerando toda a população cadastrada anualmente no CadÚnico dos dois grupos de municípios, houve um aparente impacto negativo do desastre nos primeiros dois anos, que se inverteu nos anos seguintes. Mas estas variações foram todas na terceira casa decimal de um indicador que varia de 0 a 1, ou seja, é muito arriscado afirmar que isso seja um real impacto sendo observado. Da mesma forma, o acesso ao trabalho indica pequenos impactos negativos nos três primeiros anos, enquanto a escassez de renda mostra impactos positivos em todos os períodos. Quando se isolam apenas as populações com certas características demográficas, tem-se diferenças marcantes. Dentre as famílias que possuem filhos recém-nascidos, houve aparentes impactos negativos por todos os períodos, tanto no IVF total quanto nas dimensões de trabalho e de renda. Já as famílias com a presença de idosos parecem ter sofrido impactos positivos no IVF como um todo e na dimensão de trabalho, mas efeitos negativos na escassez de renda. A falta de cônjuge mostrou grande impacto negativo nos primeiros anos que se inverte no longo prazo para o IVF global e escassez de recursos, mas a dimensão de acesso ao trabalho é prejudicada por todos os períodos. Por fim, a presença de uma minoria em idade ativa do domicílio teve resultados mais voláteis. Este trabalho não tem a pretensão de identificar o que isto significa, devido às suas limitações na base de dados e nas análises estatísticas, sendo apenas possível especular hipóteses para próximos estudos.

Palavras-Chave: Vulnerabilidade social; Pobreza Multidimensional; Desastres; Rompimento de Barragens; Desastre da Samarco/Vale/BHP; Impacto.

ABSTRACT

The collapse of the Fundão dam in Mariana (MG) demonstrates how demography can be a crucial factor in the ability of families to recover or worsen their conditions. Starting from a de-identified database of the Cadastro Único of the Ministry of Citizenship, this dissertation main objective is to show the potential that multidimensional indicators of social vulnerability have in capturing the effects of a technological disaster on the affected families, including identifying aggravating or mitigating differences in the family composition in the face of sudden and adverse changes in the conditions of access to work and scarcity of resources over the years after the disaster. Based on the case study of the Mariana disaster, an impact evaluation based on the differences in differences method was elaborated, taking as a reference its effects on the Family Vulnerability Index (IVF, in Portuguese): a multidimensional indicator of social vulnerability. An attempt was made to compare the annual results in the global indicator and in the dimensions of work and income, first by the total population of CadÚnico, and then by analyzing only the universe of families with certain demographic conditions: presence of a live-born child; presence of the elderly; absence of a spouse; and working-age minority households. The treatment group encompasses the 35 mining towns officially considered affected by the disaster. In order to list the municipalities in the control group, the K-Nearest Neighbor algorithm was used, taking 23 variables as a parameter to pair those most similar to the 35 affected, resulting in a total of 169 classified. The evaluation was divided into periods, considering each of them as a number of years from December 2014 (eleven months before the breakup), with the short-term period 2014-15; medium terms 2014-16 and 2014-17; and long term 2014-2018. This work presented apparently ambiguous results, but which manage to outline an interesting scenario to analyze and debate. Considering the entire population registered annually in CadÚnico of the two groups of municipalities, there was an apparent negative impact of the disaster in the first two years, which was reversed in the following years. But these variations were all in the third decimal place of an indicator that ranges from 0 to 1, that is, it is very risky to say that this is a real impact being observed. Likewise, access to work indicates small negative impacts in the first three years, while income scarcity shows positive impacts in all periods. When only populations with certain demographic characteristics are isolated, there are marked differences. Among the families that have newborn children, there were apparent negative impacts for all periods, both in the total IVF and in the dimensions of work and income. Families with the presence of elderly people seem to have suffered positive impacts on the IVF as a whole and on the work dimension, but negative effects on income scarcity. The lack of a spouse showed a great negative impact in the first years, which is reversed in the long term for the global IVF and scarcity of resources, but the dimension of access to work is impaired for all periods. Finally, the presence of a working-age minority in the household had more volatile results. This work does not intend to identify what this means, due to its limitations in the database and in the statistical analyses, and it is only possible to speculate hypotheses for future studies.

Keywords: Social vulnerability; Multidimensional poverty; Disasters; Dam Failure; Samarco/Vale/BHP Disaster; Impact.

LISTA DE FIGURAS

Figura 2-1 – Estimativa do número de ocupados vivendo em situação de extrema pobreza; Brasil, 2000 - 2019.....	23
Figura 2-2 – Proporção do número de ocupados em cada classe de pobreza em relação ao contingente total da força de trabalho. Brasil, 2000 - 2018.....	24
Figura 4-1 – Simulação de uma comparação entre as trajetórias do grupo de controle, do grupo de tratamento, com e sem a intervenção.....	62
Figura 4-2 – Exemplo de classificação do método K-Nearest Neighbor com $k = 1$ e $k = 3$	72
Figura 4-3 – Percentual de erro na capacidade de previsão de municípios atingidos na amostra de teste pela técnica de k-NN, por valor de K.....	78
Figura 4-4 – Níveis de sensibilidade e de especificidade na capacidade de previsão de municípios atingidos na amostra de teste pela técnica de k-NN, por valor de K.....	79
Figura 4-5 – Mapa dos municípios atingidos e dos municípios classificados como de controle..	83
Figura 5-1 – Taxas de crescimento anuais da população cadastrada no CadÚnico, entre 2012 e 2018, por grupo	85
Figura 5-2 – Índice de Vulnerabilidade das Famílias para o estado de Minas Gerais e grupos de atingidos e de controle, entre 2012 e 2018	88
Figura 5-3 – Impacto de diferenças em diferenças, por período, para as componentes da dimensão acesso ao trabalho	90
Figura 5-4 – Impacto de diferenças em diferenças, por período, para as componentes da dimensão escassez de renda	92
Figura 5-5 – Impacto no grupo de atingidos por condição de vulnerabilidade pela variável V1 (presença de filhos recém-nascidos), por períodos selecionados.	93
Figura 5-6 – Determinantes das diferenças entre o impacto sofrido pelos subgrupos de resistentes e de vulneráveis pela variável V1 (presença de filhos recém-nascidos), no período entre 2014-2018	94
Figura 5-7 – Impacto no grupo de atingidos por condição de vulnerabilidade pela variável V6 (presença de idosos com mais de 65 anos), por períodos selecionados	95
Figura 5-8 – Determinantes das diferenças entre o impacto sofrido pelo grupo total de atingidos e do subgrupo de resilientes pela variável V6 (presença de idosos com mais de 65 anos), no período entre 2014-2018.....	96
Figura 5-9 – Impacto no grupo de atingidos por condição de vulnerabilidade pela variável V7 (ausência do cônjuge do responsável pela família), por períodos selecionados.....	97
Figura 5-10 – Determinantes das diferenças entre o impacto sofrido pelo grupo total de atingidos e do subgrupo de resilientes pela variável V7 (ausência do cônjuge do responsável pela família), no período entre 2014-2018.....	97
Figura 5-11 – Impacto no grupo de atingidos por condição de vulnerabilidade pela variável V8 (minoria do domicílio em idade ativa), por períodos selecionados	98
Figura 5-12 – Determinantes das diferenças entre o impacto sofrido pelo grupo total de atingidos e do subgrupo de resilientes pela variável V8 (minoria do domicílio em idade ativa), no período entre 2014-2018.....	99

LISTA DE TABELAS

Tabela 2-1 – Escopo das análises consideradas.....	27
Tabela 2-2 – Projetos de reparação de dano executados pela Fundação Renova, conforme destacados em seu portal de notícias	32
Tabela 3-1 – Dimensões e indicadores que compõem o IPM	41
Tabela 3-2 – Variáveis componentes da dimensão Vulnerabilidade demográfica, conforme adaptadas para a base desidentificada do CadÚnico	48
Tabela 3-3 – Variáveis componentes da dimensão Acesso ao trabalho, conforme adaptadas para a base desidentificada do CadÚnico.....	49
Tabela 3-4 – Variáveis componentes da dimensão Escassez de recursos, conforme adaptadas para a base desidentificada do CadÚnico	50
Tabela 3-5 – Variáveis componentes da dimensão Acesso ao conhecimento, conforme adaptadas para a base desidentificada do CadÚnico	52
Tabela 3-6 – Variáveis componentes da dimensão Desenvolvimento infantojuvenil, conforme adaptadas para a base desidentificada do CadÚnico	53
Tabela 3-7 – Variáveis componentes da dimensão Condições de habitação, conforme adaptadas para a base desidentificada do CadÚnico	54
Tabela 4-1 – Indicadores e variáveis de acesso ao trabalho e escassez de recursos a serem avaliadas	69
Tabela 4-2 – Modelo de matriz de confusão	73
Tabela 4-3 – Relação de indicadores para a seleção de municípios para o grupo de controle através do método de aprendizagem k -NN	76
Tabela 4-4 – Percentual de erro, sensibilidade e especificidade pela técnica k -NN, por valor de K	78
Tabela 4-5 – Níveis de sensibilidade e especificidade para as cinco iterações de maior especificidade na segunda rodada de k -NN	79
Tabela 4-6 – Comparação entre o real número de atingidos e não atingidos (referência), com suas respectivas previsões k -NN, para a base de dados de teste	80
Tabela 4-7 – Níveis de sensibilidade e especificidade para as cinco iterações de maior especificidade na terceira rodada de k -NN.....	80
Tabela 4-8 – Comparação entre o real número de atingidos e não atingidos (referência), com suas respectivas previsões k -NN, para a base de dados de treinamento	81
Tabela 4-9 – Relação de municípios classificados para os grupos de atingidos e de controle.....	82
Tabela 5-1 – Nº de famílias cadastradas no CadÚnico entre 2012 e 2018, por grupo	85
Tabela 5-2 – Resultados das variáveis de parâmetro para os grupos de municípios atingidos, de controle e do estado de Minas Gerais	86
Tabela 5-3 – Resultados anuais do IVF, por grupos de atingidos e de controle; entre 2012 a 2018	87

LISTA DE SIGLAS

AASS – Amostragem aleatória simples sem reposição
BID – Banco Interamericano de Desenvolvimento
BPC – Benefício de Prestação Continuada
CadÚnico – Cadastro Único para Programas Sociais do Governo Federal
CBO – Classificação Brasileira de Ocupações
Copasa – Companhia de Saneamento de Minas Gerais
DD – Diferenças em diferenças
Emater – Empresa Brasileira de Pesquisa Agropecuária
ESF – Estratégia Saúde da Família
FJP – Fundação João Pinheiro
FGT – Foster-Greer-Thorbecke
J-PAL – Abdul Latif Jameel Poverty Action Lab
IBGE – Instituto Brasileiro de Geografia e Estatística
IDF – Índice de Desenvolvimento da Família
IDH – Índice de Desenvolvimento Humano
IPH – Índice de Pobreza Humana
IPM – Índice de Pobreza Multidimensional
IVF – Índice de Vulnerabilidade das Famílias
IMRS – Índice Mineiro de Responsabilidade Social
Ipea – Instituto De Pesquisa Econômica Aplicada.
k-NN – K-Nearest Neighbor
Loas – Lei Orgânica de Assistência Social
AF – Método Alkire-Foster
M&A – Monitoramento e avaliação
ODS – Objetivos do Desenvolvimento Sustentável
ONU – Organização das Nações Unidas
Ophi – Oxford Poverty And Human Development Initiative
OIT – Organização Internacional do Trabalho
OCDE – Organização para a Cooperação e Desenvolvimento Econômico
PPP – Paridade do poder de compra

Pnud – Programa das Nações Unidas para o Desenvolvimento

IPC – Programa de Comparação Internacional

RDD – Regressão descontínua

RDH – Relatório de Desenvolvimento Humano

Saps - Secretaria de Atenção Primária À Saude

Sedru – Secretaria de Estado de Desenvolvimento Regional, Política Urbana e Gestão Metropolitana de Minas Gerais

Sedese – Secretaria de Estado de Trabalho e Desenvolvimento Social de Minas Gerais

RCT – Seleção aleatória

Seplag - Secretaria de Estado de Planejamento e Gestão

TTAC – Termo de Transação e Ajustamento de Conduta

Unicef – Fundo Das Nações Unidas para a Infância

UFMG – Universidade Federal de Minas Gerais

Usaid – Agência dos Estados Unidos para o Desenvolvimento Internacional

VI – Variáveis instrumentais

SUMÁRIO

1	INTRODUÇÃO	16
2	MARCOS CONCEITUAIS.....	19
2.1	Pobreza multidimensional.....	20
2.2	Vulnerabilidade social	21
2.3	Trabalhadores em situação de pobreza.....	22
2.4	Determinantes demográficos para o trabalho em situação de pobreza	25
2.5	Demografia dos desastres	28
2.6	O contexto do rompimento da barragem do Fundão.....	30
3	ÍNDICES MULTIDIMENSIONAIS	34
3.1	Procedimentos metodológicos	34
3.2	Bases de dados	35
3.2.1	Cadastro Único.....	35
3.2.2	Censo 2010 e IMRS	37
3.3	Indicadores multidimensionais de pobreza e de vulnerabilidade.....	37
3.3.1	Definição	38
3.3.2	Elaboração de índices multidimensionais	39
3.4	Índice de Pobreza Multidimensional (IPM)	40
3.4.1	Método Alkire-Foster.....	41
3.5	Índice de Vulnerabilidade das Famílias (IVF)	42
3.5.1	Vulnerabilidade das famílias.....	44
3.5.2	Construção do índice.....	44
3.6	Dimensões e variáveis do IVF	45
3.6.1	Vulnerabilidade demográfica (V).....	47
3.6.2	Acesso ao trabalho (T)	48
3.6.3	Escassez de recursos (R)	49
3.6.4	Acesso ao conhecimento (C).....	51

3.6.5	Desenvolvimento infantojuvenil (D).....	52
3.6.6	Condições de habitação (H)	53
3.7	Agregação das variáveis e dimensões	54
3.7.1	Expressões para cada dimensão	55
4	AVALIAÇÃO DE IMPACTO POTENCIAL ROMPIMENTO DA BARRAGEM NA VULNERABILIDADE SOCIAL	57
4.1	Definindo avaliação de impacto	57
4.2	Avaliação de impacto com o método Diferença em Diferenças	62
4.2.1	Passo a passo para realizar uma avaliação de impacto DD:.....	64
4.3	A aplicação da avaliação de impacto neste estudo de caso.....	66
4.3.1	Comparações anuais.....	66
4.3.2	Comparações por períodos.....	67
4.3.3	Relacionamento entre dimensões demográficas e de ocupação.....	68
4.4	Seleção do grupo de tratamento.....	69
4.5	Seleção do grupo de controle.....	70
4.5.1	O algoritmo <i>K-Nearest Neighbor</i> (k-NN)	70
4.5.2	Variáveis de pareamento para o algoritmo k-NN.....	73
4.5.3	Estratégia de implementação do algoritmo <i>k-NN</i>	76
4.5.4	Municípios elencados para o Grupo de Controle	77
5	ANÁLISE DOS RESULTADOS.....	84
5.1	Descrição dos municípios dos grupos de atingidos e de controle.....	84
5.2	Impactos no Índice de Vulnerabilidade das Famílias	88
5.2.1	Diferenças em diferenças por período.....	88
5.2.2	Decomposição das dimensões de acesso ao trabalho e escassez de renda.....	89
5.3	Impactos da dinâmica demográfica nas dimensões de trabalho e de renda	92
5.3.1	Presença de filhos nascidos vivos no ano (V1)	93
5.3.2	Presença de idosos com mais de 65 anos (V6).....	95
5.3.3	Ausência do cônjuge do responsável pela família (V7).....	96
5.3.4	Minoria do domicílio em idade ativa (V8).....	98

6	CONSIDERAÇÕES FINAIS E CONCLUSÕES	100
	REFERÊNCIAS BIBLIOGRÁFICAS	105
	ANEXO I	113

1 INTRODUÇÃO

Situações adversas podem acontecer com qualquer pessoa. Uma mesma tragédia pode atingir várias famílias da mesma maneira, mas a capacidade de superar os problemas causados pelo imponderado pode ser muito diferente tanto nas decisões e estratégias quanto nos recursos e nas capacidades. Neste sentido, o interesse público deve se situar não apenas nos riscos de desastres por si só, mas com as maneiras que cada família possui para enfrentá-los a fim de mitigar seus prejuízos para a sociedade: dentre eles está o potencial de agravar suas condições de pobreza.

Diante dos dois casos graves de desastres tecnológicos em Minas Gerais – o rompimento da barragem de Fundão de Mariana em 2015 e da barragem da Mina do Feijão de Brumadinho em 2019 – a relação entre riscos ambientais e seus impactos na sociedade foi alvo de muitos estudos no Brasil, especialmente inseridos em um cenário acerca dos problemas causados pelas mudanças climáticas (REZENDE, 2021). Trata-se de dois estudos de casos relevantes para entender como prevenir e remediar tragédias futuras, além de alertar a sociedade como um todo de que vidas estão sob risco se a omissão prevalecer.

A evolução da atuação do Estado, do terceiro setor e das organizações multilaterais no combate à pobreza tem mostrado que o próprio conceito de “pobreza” é bem mais complexo do que a simples escassez de renda. Por exemplo, a transição demográfica ocorreu em países que não necessariamente enriqueceram em termos de renda e produção: pessoas passaram a conviver com melhores condições de saúde e aumentar a esperança de vida mesmo em condições precárias de infraestrutura, emprego e renda.

Assim, começou a surgir a noção de que a pobreza é fundamentalmente multidimensional, ou seja, induzida por fatores e características que vão além do crescimento econômico clássico. Esta nova forma de analisar e combater a pobreza ainda é disputada tanto na esfera científica quanto política, como será mostrado neste trabalho, ainda que tenha se tornado cada vez mais reconhecida.

Um dos aspectos mais importantes dessa identificação multidimensional é o fenômeno das pessoas ocupadas em situação de pobreza: aquelas que, mesmo conseguindo emprego e fontes de renda, ainda não conseguem superar as condições precárias em que vivem.

Pessoas em situação de vulnerabilidade podem permanecer ocupadas, mas com redução na margem extensiva do trabalho ou aumento da precarização dos postos de trabalho após a ocorrência de desastres. O nível salarial também pode ser afetado de forma significativa, afetando a dimensão

monetária da pobreza. Alteração na composição dos domicílios também representar uma resposta ou uma consequência direta desses desastres.

Diante deste contexto, este trabalho procura avaliar um exemplo de como esta conexão entre riscos ambientais e pobreza pode ser estruturada. Procura-se analisar as **composições familiares** de pessoas em situação de pobreza. A pobreza é compartilhada por um núcleo familiar¹, e quando efeitos adversos ocorrem com um de seus membros, as características demográficas dos demais podem se tornar o fiel da balança entre as oportunidades de sair ou aprofundar ainda mais a pobreza. O método de mensuração que será apresentado é o uso de índices multidimensionais, escaláveis e decomponíveis por cada potencial fator associável de vulnerabilidade. Foi utilizado o Índice de Vulnerabilidade das Famílias (IVF), composto por seis dimensões de vulnerabilidade, dentre elas as composições demográficas e as variáveis de trabalho e de renda. Pela sua simplicidade de cálculo e possibilidade de ser decomposto por dimensão, é possível avaliar as relações determinantes entre família e acesso ao trabalho e a recursos.

Este é um estudo de caso que avalia o impacto do rompimento da barragem de Fundão no Índice de Vulnerabilidade das Famílias dos municípios atingidos a partir da ótica das famílias (sua composição) e das condições gerais (status e renda) das pessoas ocupadas em situação de pobreza.

Esta dissertação está organizada em cinco capítulos, além desta introdução. O segundo capítulo contém o marco conceitual que fundamenta tecnicamente os conceitos citados nesta introdução (pobreza multidimensional, pessoas ocupadas em situação de pobreza, desastres tecnológicos e vulnerabilidade), e situa sua relevância no debate político. O capítulo também introduz o contexto do rompimento da barragem de Fundão, o estudo de caso referido neste trabalho.

O terceiro capítulo apresenta em mais detalhes as principais características dos indicadores multidimensionais, caracterizando em especial o IVF, as bases de dados utilizadas, suas dimensões e variáveis, bem como seu potencial para um estudo de caso usando dados nacionais. O IVF e suas componentes foram calculados em níveis médios para toda a população disponível nas bases anuais desidentificadas do Cadastro Único, anualmente entre 2012 e 2018.

¹ Sem deixar de reconhecer que mesmo dentro de uma estrutura familiar há desigualdades, inclusive de renda: assume-se que este trabalho é incapaz de tocar em um assunto ainda mais complexo e de difícil identificação.

O quarto capítulo, por sua vez, descreve a estratégia de avaliação de impacto adotada para este estudo: uma avaliação por diferenças em diferenças, entre a variação dos resultados dos IVFs anuais para dois grupos de municípios: os atingidos pelo desastre e um grupo de controle, identificado por pareamento. Trata-se de um conjunto de municípios cujos determinantes de variações no IVF são bastante similares àquelas dos atingidos, a ponto de podermos assumir que as diferenças de trajetória dos dois grupos nos ajudem a identificar os efeitos do rompimento em vários períodos. Considerando que o rompimento ocorreu em 05 de novembro de 2015, os efeitos foram comparados a partir da evolução do indicador nos dois grupos com início em 2014, analisando a variação de curto prazo (2014 a 2015), de médio prazo (2014 a 2016 e 2014 a 2017) e de longo prazo (2014 a 2018). O método de seleção destes municípios de controle está também descrito neste capítulo.

O quinto capítulo traz a análise dos resultados da avaliação de impacto. Ele começa com uma descrição dos municípios envolvidos nos dois grupos (tratamento – atingidos – e controle – pareados); em seguida, traz a avaliação dos impactos no IVF para os dois grupos, bem como em suas seis dimensões, detalhando até o nível das variáveis para os casos do acesso ao trabalho e escassez de recursos. Após a avaliação geral, foram realizadas avaliações para as populações específicas dentro dos dois grupos, dividindo-os a partir de quatro composições demográficas diferentes (presença de crianças, de idosos, de cônjuges e de dependentes), especificamente nas dimensões de vulnerabilidade pelo acesso ao trabalho e pela escassez de renda. Na medida em que os resultados empíricos são demonstrados, propostas de hipóteses para se aprofundar no tema e encontrar melhores relações de causas são indicadas.

O último capítulo conclui a avaliação, com foco em descrever o potencial destas ferramentas para avaliação de impactos multidimensionais resultantes de desastres tecnológicos. As informações geradas por esta avaliação são importantes por si só, mas merecem a atenção também pelo seu potencial de gerar hipóteses para identificar determinantes demográficos dos diferenciais de vulnerabilidade do trabalho em condição de pobreza.

2 MARCOS CONCEITUAIS

O combate à pobreza ainda está longe de se esgotar. Como ressalta Pinker (2018):

“A pobreza também dispensa explicações. Em um mundo governado pela entropia e evolução, é o estado-padrão da humanidade. A matéria não se arranja espontaneamente para ser abrigo ou roupa, e seres vivos fazem tudo para não se tornar nossa comida. Como observou Adam Smith, o que precisa ser explicado é a riqueza.”

A partir da revolução industrial, o progresso econômico e humano atingiu patamares incrivelmente mais elevados do que em toda a sua história. Mas também é igualmente claro que esta nova riqueza não foi distribuída de forma minimamente justa pela população. O resultado dessa desigualdade distributiva da riqueza gerada reflete-se na persistência de um contingente populacional em condições de pobreza e miséria, com maior ou menor intensidade ao redor do globo. Ainda que tenha havido progresso na redução do número de pobres, a vulnerabilidade econômica (especialmente monetária e alimentar) continua a assolar um enorme contingente populacional. Entre os anos de 1990 e 2010, o globo observou sua maior redução da pobreza da história: cerca de um bilhão de pessoas deixaram a condição de extrema pobreza, reduzindo-a de 41% para 21% em apenas duas décadas (THE ECONOMIST, 2013). Ainda assim, atualmente cerca de 8% da população global continua a viver em condição extremamente pobre (WDL, 2022).

O objetivo global e foco no combate à pobreza é refletido nas diretrizes das Nações Unidas. O primeiro dos Objetivos do Desenvolvimento Sustentável (ODS) estabelecido pelas Nações Unidas para a Agenda 2030 é "acabar com a pobreza em todas as suas formas, em todos os lugares" (ONU, [2020]), demonstrando um comprometimento, ou pelo menos um reconhecimento, de todos os países na organização de agir para cumprir este objetivo. Especificamente, uma das metas incluídas neste objetivo é:

"1.5 Até 2030, construir a resiliência dos pobres e daqueles em situação de vulnerabilidade, e reduzir a exposição e vulnerabilidade destes a eventos extremos relacionados com o clima e outros choques e desastres econômicos, sociais e ambientais" (ONU, [2020]).

Logo, a preocupação com a erradicação da pobreza vai além da simples elevação da renda dos indivíduos, mas alterar sua condição em todas as suas dimensões, promovendo a capacidade destas pessoas de resistir a desafios que podem ameaçá-las a retornar à pobreza. A própria ONU,

usando este conceito mais complexo de pobreza, estima que o número de pessoas em condição de pobreza multidimensional seja de 1,3 bilhões (PNUD, 2019).

Chakravarty e D'Ambrosio (2013) sugerem que, para que a formulação de políticas públicas que visem atingir tal meta, os planejadores sociais precisam identificar as causas da pobreza de modo a traçar estratégias específicas que sejam efetivas justamente em eliminar estes problemas, identificando as distintas características dos indivíduos. Eles argumentam que também são necessárias condições de se quantificar o tamanho do problema, para então saber o que priorizar diante de recursos escassos.

Neste capítulo, faz-se uma revisão acerca dos principais conceitos envolvidos. Inicialmente são descritos os conceitos ligados às condições socioeconômicas dos indivíduos e das famílias: pobreza multidimensional, vulnerabilidade social, o trabalhador em situação de pobreza e os fatores demográficos determinantes desta condição. Em seguida tem-se uma breve explicação sobre a demografia dos desastres, bem como uma contextualização do rompimento da barragem de Fundão.

2.1 Pobreza multidimensional

Chakravarty e D'Ambrósio (2013) fazem uma extensa discussão sobre a necessidade de se definir pobreza não apenas como fonte de renda limitada, mas de forma multidimensional. A identificação da situação de pobreza pode ser baseada na comparação da renda com uma linha predefinida: uma pessoa com rendimentos menores que esta linha seria identificada como pobre. Mas a definição desta linha é, de certa forma, controversa: por exemplo, governos diferentes acabam adotando cálculos próprios para estabelecê-la.

O Banco Mundial (2015) passou a adotar o valor de 1,9 dólares por dia, utilizando a conversão cambial por paridade de poder de compra, conforme mensurada em 2011. Relatórios mais recentes do Banco Mundial (2018) adotaram linhas com valores distintos para países de renda média-baixa (USD 3,2) ou média-alta (USD 5,5).

No Brasil, a lei mais recente a tratar do tema (Lei Federal n.º 14.284/2021), define como domicílio extremamente pobre aquele cuja renda per capita mensal é de até R\$ 105,00, e pobre aquele que com renda per capita de R\$ 105,01 a R\$ 210,00 (BRASIL, 2021). Por outro lado, a Lei Orgânica de Assistência Social (Loas) direciona o Benefício de Prestação Continuada (BPC) ao

idoso que "não possuir meios de prover a própria manutenção nem de tê-la provida por sua família", e que se enquadra nesta condição aqueles que possuam uma renda per capita de 1/4 do salário-mínimo vigente (BRASIL, 1993).

Além de uma linha de corte de renda, a composição da pobreza envolveria a identificação de privações que as pessoas sofram em outras dimensões e que dificultam suas chances prosperar ao longo da vida (SEN, 1975). Alfabetização, moradia, saúde, provisão de bens públicos, alimentação adequada, serviços essenciais e comunhão com amigos são exemplos de dimensões do bem-estar relacionadas às capacitações (*capabilities*). Nesse sentido, a preocupação é com o problema da insuficiência dessas capacidades. Portanto, seria preciso investigar este prisma de múltiplas dimensões potenciais causadoras da situação de pobreza (Chakravarty e D'Ambrosio, 2013).

2.2 Vulnerabilidade social

Em linhas gerais, a vulnerabilidade pode ser definida como uma medida de exposição à possibilidade de se sofrer algum prejuízo. No caso da vulnerabilidade social, como um termo mais técnico que visa ser utilizado em estudos ou políticas, Monteiro (2011) propõe a criação de um marco conceitual, com fundamentos e determinantes mais objetivos, permitindo uma melhor sinergia entre analistas.

No dicionário sobre trabalho, profissão e condição docente elaborado pela Universidade Federal de Minas Gerais (UFMG), Ximenes (2010) define vulnerabilidade social como "um conceito multidimensional que se refere à condição de indivíduos ou grupos em situação de fragilidade, que os tornam expostos a riscos e a níveis significativos de desagregação social". Trata-se da insegurança de se sofrer consequências econômicas ou sociais devido à falta de capacidades de enfrentar cenários agravantes, como a destruição da barragem que será analisada neste estudo (XIMENES, 2010).

Marcadamente a partir dos anos 1990, as definições mais clássicas de pobreza se mostraram limitadas nos debates para o enfrentamento de problemas além da falta de recursos econômicos, como no surgimento do combate à contaminação mundial pelo HIV (MONTEIRO, 2011). Protagonizada por organismos multilaterais, houve uma difusão de debates públicos acerca da vulnerabilidade social, preocupando-se com riscos que não se apresentassem apenas para a

população já empobrecida, mas também para a desproteção da sociedade como um todo. Mesmo famílias que não sejam pobres, se incapazes de lidar com circunstâncias negativas podem sofrer consequências graves, inclusive a de se tornarem pobres.

Em sua análise, Monteiro (2011) considera que a vulnerabilidade social não é, necessariamente, uma condição inerente, mas uma predisposição ao agravamento da situação social do vulnerável. É uma exposição a riscos sociais, econômicos, culturais ou ambientais, que podem ser enfrentadas e mitigadas. A autora identifica que um ponto crucial para o conceito de vulnerabilidade é sua relação com transições e com a ocorrência de efeitos indesejados. A resiliência da pessoa contra as adversidades que surjam pode ser analisada a partir dos ativos que ela possui, aqui definidos como "um conjunto articulado de condições que irão implicar a qualidade, quantidade e diversidade dos recursos internos, os quais devem ser avaliados a partir de quatro aspectos: físico, financeiro, humano e social" (MONTEIRO, 2011, p. 34).

A vulnerabilidade social, quando associada a questões de natureza ambiental, assume uma dimensão ainda mais complexa. Gomes e Guedes (2021), por exemplo, fazem um estudo sobre vulnerabilidade socioambiental, argumentando que se trata de uma definição sem consenso definido. Os autores elencam três características principais na conceituação de vulnerabilidade socioambiental: a exposição ao risco (uma medida de sensibilidade), a capacidade adaptativa e a resiliência.

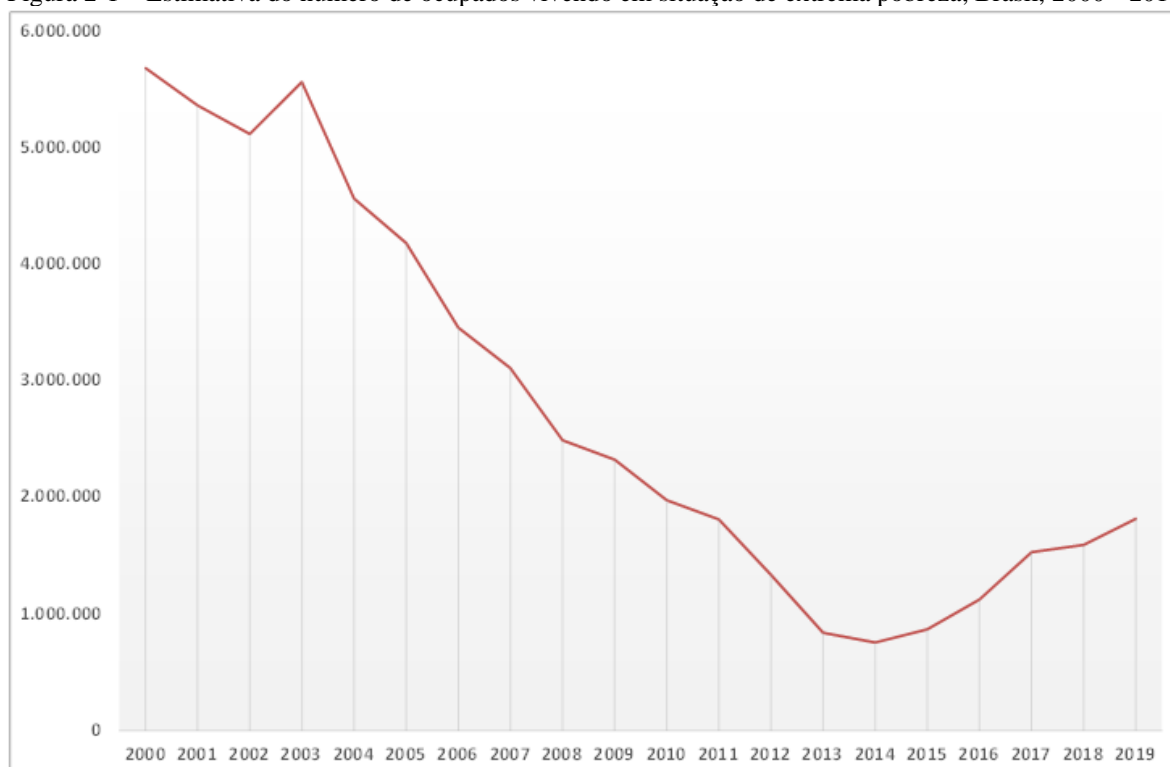
2.3 Trabalhadores em situação de pobreza

Uma condição de pobreza e vulnerabilidade marcante que tem ganhado destaque no debate público é aquela em que a pessoa tem uma ocupação como fonte de renda e, mesmo assim, vive em uma família considerada pobre. Ou seja, quando o trabalho não é suficiente para superar essa situação.

A Organização Internacional do Trabalho (OIT) considera em situação de pobreza¹ o ocupado de 15 anos ou mais de idade, cuja renda domiciliar per capita seja de USD 3,20/dia a preços internacionais de 2011. Utiliza-se a paridade do poder de compra (PPP) de 2011, segundo o Programa de Comparação Internacional (ICP) do Banco Mundial (OIT, 2018). Quando a renda é até a linha de USD 1,90/dia, considera-se o trabalhador extremamente pobre.

Com base nessas definições, estima-se que em 2018 13% dos trabalhadores no mundo eram moderadamente pobres e 8% eram extremamente pobres (GAMMARANO, 2019). No mesmo período, o Brasil possuía 2,2% e 1,8% dos seus trabalhadores como pobres e extremamente pobres, respectivamente. Mesmo sendo um dos países de renda média, e um dos maiores produtores mundiais de riqueza e alimentos, sua força de trabalho sob condição de pobreza é de mais de 3,5 milhões de ocupados (Ilostat, 2019).

Figura 2-1 – Estimativa do número de ocupados vivendo em situação de extrema pobreza; Brasil, 2000 - 2019



Fonte: Ilostat.

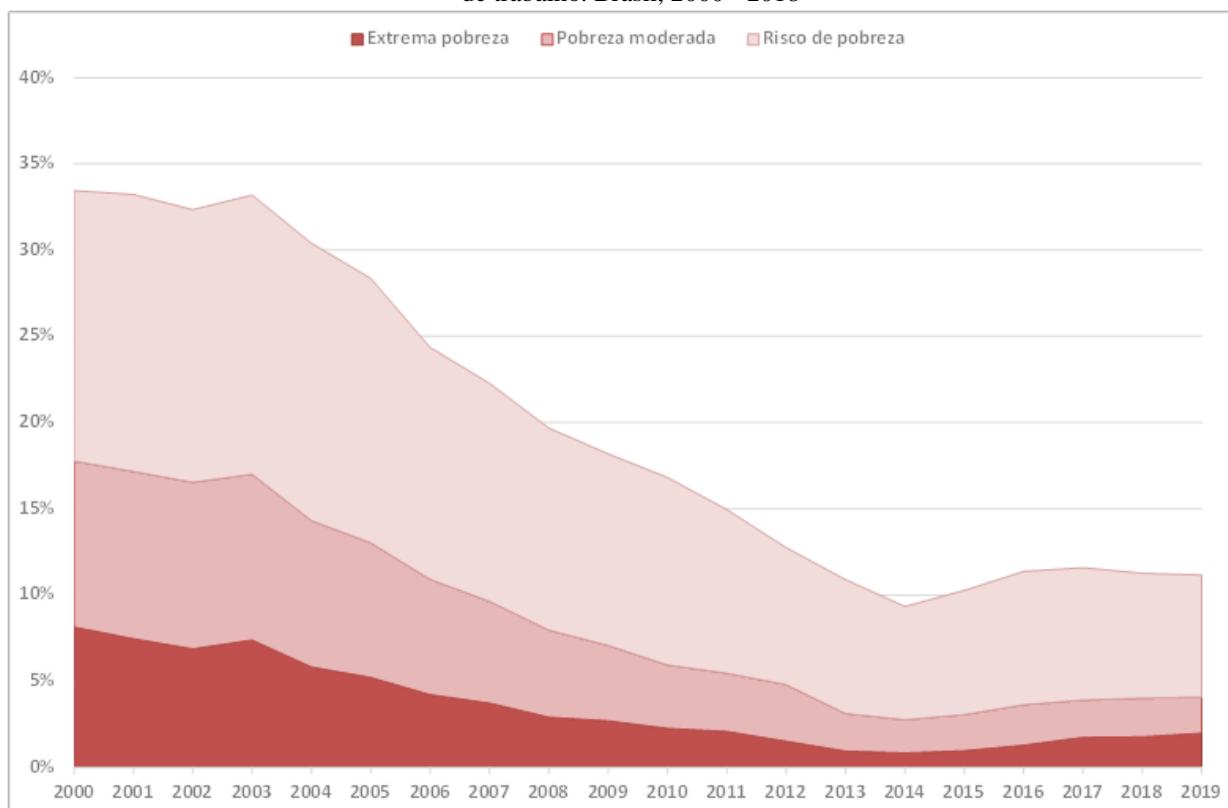
Notas: Última atualização: 26/01/2022; Link para reprodução dos dados: <https://bit.ly/3apspxl>

A Figura 2-1 representa o número estimado dos trabalhadores brasileiros em situação de pobreza extrema, ao longo do século XXI (Ilostat, 2019). Observa-se que houve um declínio acentuado ao longo das décadas de 1990 e 2000, mas que passou por um ponto de inflexão em 2014 e, desde então, esse contingente tem aumento ano após ano.

É possível também ver situação preocupante nas linhas de pobreza moderada e no risco de se entrar na condição de pobreza. A Figura 2-2 mostra o percentual de ocupados em cada faixa, em

comparação com o total da força de trabalho estimada para o Brasil entre 2000 e 2019. Estes dados são evidência de que estar empregado não é condição suficiente para superar a pobreza.

Figura 2-2 – Proporção do número de ocupados em cada classe de pobreza em relação ao contingente total da força de trabalho. Brasil, 2000 - 2018



Notas: Fonte: Ilostat. Última atualização: 26/01/2022. Link para reprodução dos dados: <https://bit.ly/3apspxl>

Ao elaborar um estudo para a OIT descrevendo estimativas de ocupação por classe econômica, Grammano (2019) pontua que no mundo todo houve uma tendência de redução dos ocupados empobrecidos em 30 anos, mas que passaram por um ponto de inflexão nos últimos anos. Ela especifica que os trabalhadores mais jovens (de 15 a 24 anos) são mais prevalentes em condições de trabalho e de pobreza. O fator mais natural de se considerar são os salários e os rendimentos reduzidos destes ocupados. Outros fatores podem ser investigados, como a quantidade de horas trabalhadas ou a alta correlação com a informalidade. Mais interessante para esta dissertação, no entanto, são os fatores relacionados à composição domiciliar:

"para tirar conclusões precisas sobre as características e determinantes da pobreza, é importante estudar a taxa de pobreza no trabalho juntamente com outros indicadores do

mercado de trabalho e com **informações sobre a composição e estrutura dos agregados familiares**" (tradução e grifo próprios; GAMMARANO, 2019).²

Apesar do relatório de Gammarano (2019) ser totalmente baseado na definição estipulada pela OIT, assim como no caso do estabelecimento de linhas de pobreza, ou mesmo na concepção de vulnerabilidade social, governos e organizações diferentes adotam critérios próprios para esta condição, especialmente para a aplicação de políticas de assistência social.

Há outras definições oficiais sobre ocupados em situação de pobreza. Por exemplo, a Organização para a Cooperação e Desenvolvimento Econômico (OCDE) estabelece que um ocupado é pobre se convive em um domicílio pobre; esta linha de pobreza domiciliar é calculada com base na metade da mediana da renda disponível em determinado país (OCDE, 2022). Já a Eurostat (2014) utiliza como base "pessoas na população total que declararam estar a trabalhar (por conta de outrem ou por conta própria) que estão em risco de pobreza, ou seja, com rendimento inferior ao limiar do risco de pobreza, fixado em 60% do rendimento disponível médio equivalente nacional (após transferências sociais)", em tradução livre.

2.4 Determinantes demográficos para o trabalho em situação de pobreza

Um ponto em comum destas definições é que se trata de uma combinação entre status individual e características domiciliares. Trata-se de atribuir a um indivíduo, que tenha uma ocupação gerando renda, uma condição que depende não só desta fonte, mas do somatório dos recursos adquiridos pelos membros do domicílio, independentemente da origem e de sua distribuição entre todos. Neste sentido, toda discussão sobre trabalhadores pobres passa por características adquiridas pela sua composição domiciliar: membros dependentes, dedicação às tarefas do lar, investimentos em educação de filhos, situação conjugal, morbidade, dentre outros.

Visando aprofundar a discussão sobre essa relação intrínseca entre a demografia e a pobreza da população ocupada, os autores Polizzi, Struffolino e Van Winkle (2022) elaboraram uma meta-

2 Os termos originais em inglês foram traduzidos livremente neste trabalho. "Extremely poor" foi traduzido como extremamente pobre, "moderately poor" como moderadamente pobre, e "near poor" como em risco de pobreza; apesar de que, neste último caso, a tradução não fora literal, a intenção é manter clareza no texto. Além disso, assumiu-se como pobre o ocupado cuja condição seja extremamente ou moderadamente pobre.

análise bastante abrangente acerca de estudos que buscam encontrar determinantes entre características demográficas específicas e os riscos de se tornar um ocupado pobre. Nesse sentido, eles encorajam a produção de análises, de preferência, longitudinais, avaliando a relação entre processos demográficos familiares e a pobreza no trabalho.

"Os estudos em sociologia e economia referentes à pobreza no trabalho se concentraram, até agora, principalmente no crescimento dos setores com baixos salários e na seleção do indivíduo por ocupações de baixa remuneração, ao invés da pobreza no trabalho como um fenômeno mais amplo" (POLIZZI, STRUFFOLINO e VAN WINKLE, 2022).

Polizzi, Struffolino e Van Winkle (2022) descrevem em sua meta-análise seus critérios para procurar por publicações sobre o tema, filtrar aquelas que não correspondem com seus critérios, e separar em unidades de análise os estudos encontrados (ou seja, uma única publicação pode conter mais de uma análise). Eles encontram 898 publicações ao todo, mas após seus filtros, apenas 30 foram considerados. Ao todo, foram utilizadas 84 unidades de análises empíricas. A Tabela 2-1 resume características das análises identificadas.

Alertando que a existência de trabalhadores em situação de pobreza pode causar exclusão social mesmo em democracias bem estabelecidas, o foco dos autores foi em pesquisas empíricas com dados de países europeus e da OCDE, em que buscaram compreender o efeito estatístico de quatro dimensões demográficas na situação de pobreza dos ocupados: transição para a vida adulta, uniões estáveis, fecundidade e dissoluções de casais (POLIZZI, STRUFFOLINO e VAN WINKLE, 2022). Embora os autores considerem que sua revisão seja capaz de identificar as causas que fortalecem a relação entre demografia e pobreza no trabalho, eles consideram que há evidências apoiando as principais hipóteses acerca destas ligações.

Um ponto relevante para a meta-análise fora considerar as diferentes definições de trabalho em situação de pobreza utilizadas nos estudos originais. A Comissão Europeia define como "*working poor*" os adultos em idade ativa (18 a 64 anos) ocupados em pelo menos sete dos últimos 12 meses e vivendo em um domicílio cuja renda per capita seja menor que 60% da renda domiciliar per capita mediana de seu país².

Uma diferença importante entre a OCDE e a Comissão Europeia é que, enquanto a Comissão Europeia conta o indivíduo ocupado como o "*working poor*", a OCDE estende esse adjetivo a todos os membros do domicílio, o que exige muita cautela nas comparações.

Para o governo federal dos EUA, a definição oficial abrange ocupados em metade das semanas do último ano, de 15 anos ou mais e vivendo abaixo da linha oficial de pobreza estabelecida.

Tabela 2-1 – Escopo das análises consideradas

Escopo	Nº de análises	% dos 84
Natalidade	76	90%
União estável	57	68%
Casamento	9	11%
Dissolução da união estável	11	13%
Domicílios como unidades de análise	6	7%
Transversais	72	86%
Longitudinais	12	14%

Fonte: Polizzi, Struffolino e Van Winkle (2022).

A crise financeira global de 2008 precisa ser considerada, porque não só causou graves problemas de desemprego, como também no empobrecimento das famílias com provedores de renda ocupados (POLIZZI, STRUFFOLINO e VAN WINKLE, 2022). O comportamento das famílias diante de crises se altera, por exemplo, adiando casamentos e filhos, além de agravar a situação daqueles que passavam por alguma transição no ciclo de vida justamente neste período. Além disso, os fenômenos demográficos de longo prazo também podem indicar direções sobre o futuro das condições dos ocupados, como a queda da fecundidade, o adiamento da saída do domicílio dos pais para a formação de um próprio e as uniões estáveis, eventos que ocorrem enquanto a carreira profissional dos indivíduos também se desenvolve.

Polizzi, Struffolino e Van Winkle (2022) consideram que existem três classes de marcos teóricos que podem conectar a dinâmica familiar com problemas sociais. O primeiro seria a seletividade, que se relaciona com causalidades reversas, por exemplo, com as escolhas divergentes entre homens e mulheres dadas as condições do mercado de trabalho para cada sexo. A causalidade, por sua vez, analisa as relações entre mudanças que ocorrem na dinâmica demográfica e seus efeitos nas condições de ocupação e pobreza (por exemplo, a penalidade à maternidade no mercado de trabalho). Por fim, as análises mecanicistas estudam as mudanças nas próprias demandas dos domicílios ao longo do seu ciclo de vida, como ocorre nos estudos da demografia econômica.

Dos resultados estatisticamente relevantes na literatura encontrados por Polizzi, Struffolino e Van Winkle (2022), cerca de 40% (25 em 66) das análises encontraram relação estatisticamente

significativa entre a formação de uniões e casamentos e a redução da condição de ocupado pobre. Por outro lado, tanto a natalidade quanto a dissolução das uniões provocaram aumentos nas chances de empobrecimento dos ocupados em metade das análises de cada tema. Houve mais resultados corroborando estas conclusões, mas não foram considerados estatisticamente significantes.

Os autores também identificaram efeitos institucionais e de políticas de assistência social na relação entre dinâmicas demográficas e risco de pobreza no trabalho. O caso mais destacado é o de que certos Estados de bem-estar social conseguiram demonstrar relação negativa significativa entre natalidade e trabalhadores pobres, principalmente em países da Europa central e oriental, inclusive nos países bálticos.

2.5 Demografia dos desastres

A relação entre demografia e grandes choques na população sempre fez parte desta ciência. A demografia histórica, por exemplo, possui diversos estudos sobre os efeitos das crises de mortalidade, causadas por epidemias de doenças infecciosas ou desastres naturais (LIVI-BACCI, 1984). Mais recentemente, temas relacionando impactos nas dinâmicas demográficas causados pelas mudanças estruturais contemporâneas vêm ganhando força no debate mundial. Por exemplo, os impactos demográficos do aumento de crises e de desastres causados pelas mudanças climáticas (HORWOOD E REITANO, 2016); as causas e os efeitos provocados por deslocamentos forçados, tanto nas regiões de origem como nos destinos (PNUD, 2016); conflitos armados e seus danos para além da destruição inicial, como o agravamento mortalidade dos inocentes, seja por mortes diretas ou indiretas, inclusive os efeitos danosos da deterioração dos serviços de saúde para as pessoas afetadas (BRUNBORG E URDAL, 2005).

A demografia dos desastres é a expressão que se tem utilizado para descrever os efeitos de desastres socioambientais causados pela destruição de grandes obras de infraestrutura na dinâmica demográfica das populações atingidas (KARÁCSONYI, TAYLOR E BIRD, 2020). São exemplos de temas abordados: migrações pressionadas pela deterioração das condições econômicas locais; problemas de saúde pública, na medida em que aumentam os riscos de maiores mortalidade e morbidade das populações; alterações de comportamento e composição das famílias, tais como mudanças na fecundidade e na nupcialidade das famílias, uma vez que perspectivas e preocupações

socioeconômicas dos domicílios também podem ser afetadas (BRUNBORG E URDAL, 2005; KARÁCSONYI, TAYLOR E BIRD, 2020; REZENDE, 2021).

Segundo Bowker e Chambers (2015), houve um aumento na frequência de incidentes muito graves nas infraestruturas de extração de minério, na medida em que houve um crescimento global da demanda e uma necessidade de se extrair estes recursos em lugares cada vez mais complexos de serem alcançados³. Seus danos nas populações atingidas, embora inquestionáveis, precisam ser efetivamente mensurados (REZENDE, 2021). O aumento dos estudos de demografia aplicada a desastres busca em parte contribuir para a criação de novas e melhores políticas públicas de mitigação de riscos através de estudos de casos, estimações, propostas de marcos teóricos, dentre outras contribuições (KARÁCSONYI, TAYLOR E BIRD, 2020).

Cabe destacar que há um legado cultural e institucional após a tragédia: principalmente quando comunidades, governos e empresas envolvidos conseguem chegar em acordos para a criação de investimentos em políticas de mitigação dos danos: o que pode influenciar as mudanças para além da destruição inicial (BIRD e TAYLOR, 2021). Além disso, como alerta Alvalá e Barbieri (2017), uma má administração e investimentos precários em prevenção de desastres podem efetivamente deixar as populações ainda mais vulneráveis, dado que estes eventos podem induzir maiores desigualdades nas populações atingidas.

É preciso também entender que variáveis demográficas tendem a ser mensuradas no longo prazo, em processos que ocorrem no curso de vida dos envolvidos. Além disso, este tipo de destruição tende a ser localizada em regiões rurais ou urbanas com poucos habitantes, em que geralmente é muito difícil de se encontrar dados para estudar sua dinâmica demográfica. Por exemplo: Bird e Taylor (2021) puderam discutir as variáveis demográficas antes e depois de um incêndio em uma mina de carvão na cidade de Hazelwood, na Austrália, graças aos censos australianos ocorridos em anos como 2011, 2014 e 2016, permitindo estimativas em nível municipal; já Rezende (2021) precisou recorrer aos registros administrativos do Censo Escolar para

³ Os autores definem um incidente muito grave (“*very serious*” *no original*) como provocando uma liberação de pelo menos 1 milhão de metros cúbicos e/ou percorrendo um trajeto de pelo menos 20 km e/ou causando cerca de 20 mortes ou mais (BOWKER E CHAMBERS, 2015).

estimar os impactos de migração após o rompimento da barragem de Fundão, por não haver melhores dados representativos⁴.

2.6 O contexto do rompimento da barragem do Fundão

Dada a sua relevância econômica, ambiental, social e humana, o rompimento da barragem do Fundão tem sido extensamente estudado, principalmente por autores brasileiros, sob as mais diversas perspectivas: saúde, meio ambiente, economia, direito, geografia, seus efeitos imediatos e de longo prazo. Já existe um grande legado de trabalhos acadêmicos contando em detalhes o que ocorreu. Rezende (2021), especificamente, considera que o desastre deveria ser estudado inclusive a partir da perspectiva da demografia dos desastres – um campo em crescente expansão, que situa o desastre a partir de suas consequências sobre as populações humanas em função de sua composição por sexo, idade, cor/raça e composição domiciliar, bem como da capacidade dessas características atenuarem ou agravarem as consequências.

O rompimento da barragem de Fundão, então administrada pela empresa Samarco S.A., ocorreu no dia 5 de novembro de 2015. Oficialmente, registraram-se 19 mortes oficiais, mais uma reconhecida apenas pela população local (BRASIL, 2022b).

"O colapso da estrutura da barragem do Fundão ocasionou o extravasamento imediato de aproximadamente 40 milhões de metros cúbicos de rejeitos de minério de ferro e sílica, entre outros particulados, outros 16 milhões de metros cúbicos continuaram escoando lentamente. O material liberado logo após o rompimento formou uma grande onda de rejeitos [...].

Após percorrer 22 km no rio do Carmo, a onda de rejeitos alcançou o rio Doce, deslocando-se pelo seu leito até desaguar no Oceano Atlântico, no dia 21 de novembro de 2015, no distrito de Regência, no município de Linhares (ES)." (BRASIL, 2022b)

Para um estudo descritivo detalhado e compreensivo dos eventos, existem três fontes importantes de informações oficiais.

A primeira delas foi um diagnóstico realizado em fevereiro de 2016 pelo Grupo Força-Tarefa, conforme o Decreto Estadual nº 46.892/2015 do governo do estado de Minas Gerais,

⁴ Ressaltando que o Censo Demográfico brasileiro, planejado inicialmente para ocorrer em 2020, foi adiado para agosto de 2022 (IBGE, 2022). Portanto, os últimos dados disponíveis de pesquisas domésticas municipais são de 2010.

organizado pela Secretaria de Estado de Desenvolvimento Regional, Política Urbana e Gestão Metropolitana de Minas Gerais (Sedru). Este diagnóstico descreve os danos ambientais, materiais e humanos, primeiro em escala microrregional, contendo os municípios atingidos pela lama, e depois em escala macrorregional, contendo todos os considerados atingidos de alguma forma pelos rastros deixados após o rompimento. Suas fontes de dados são principalmente estatais, a partir de diversos órgãos distintos, como a Companhia de Saneamento de Minas Gerais (Copasa), a Empresa Brasileira de Pesquisa Agropecuária (Emater) e a então Secretaria de Estado de Trabalho e Desenvolvimento Social de Minas Gerais (Sedese). O relatório contextualiza e indica as primeiras medidas, bem como produz novas recomendações de ações de mitigação por parte da Samarco S.A., responsabilizada diretamente pelo incidente (MINAS GERAIS, 2016).

A outra fonte oficial é o portal do Caso Samarco, do Ministério Público Federal (BRASIL, 2022b), que detalha uma linha do tempo de todos os principais eventos e iniciativas ocorridas após o rompimento da barragem. Nele há cópias de todos os acordos assinados, principais relatórios e pareceres técnicos, além de notícias atualizadas sobre o andamento do caso.

Após o desastre, as empresas envolvidas, Samarco S.A. e suas controladoras Vale S.A. e BHP Billinton LTDA.), comprometeram-se formalmente em criar a Fundação Renova a partir de um acordo com os órgãos dos governos de Minas Gerais, do Espírito Santo e da União. Seu objetivo é promover projetos de mitigação dos danos causados pelo desastre. Ela se comprometeu formalmente em atuar a partir de 42 projetos a serem implementados em cerca de 670 quilômetros de área impactada. A Fundação atualmente possui um portal na internet contendo prestações de contas e notícias atualizadas sobre a sua atuação (FUNDAÇÃO RENOVA, 2022).

Após o evento, mudanças estruturais fundamentais vêm ocorrendo nas instituições locais. Os governos destes municípios passaram a ter forte atuação conjunta, com vistas a ganhar maior poder de barganha das mesas de negociação. Órgãos federais e estaduais, como as secretarias de estados e os ministérios públicos, têm atuado bastante para regular e intermediar a atuação da fundação, além de eles próprios promoverem projetos para reforçar políticas públicas locais (BRASIL, 2022B).

A Tabela 2-2 lista os programas, conforme noticiados pela própria Fundação Renova em seu portal, considerando sua área de atuação e escopo. Não há espaço aqui para avaliar estes programas: o importante é ressaltar que tais projetos têm o potencial de impactar a vulnerabilidade das famílias através, principalmente, da mitigação danos (processo reparatório). Por isso a

importância de considerá-los como parte do efeito da destruição, uma vez que são de fato defesas institucionais contra ela.

Tabela 2-2 – Projetos de reparação de dano executados pela Fundação Renova, conforme destacados em seu portal de notícias

Área	Programa	Escopo
Assistência social	Programa de Apoio à Saúde Física e Mental da População Impactada	Prestação de apoio técnico às prefeituras na execução dos planos de ação de saúde.
Infraestrutura	Reconstrução de Vilas	Construção de novas localidades que irão abrigar as comunidades de Bento Rodrigues, Paracatu de Baixo e Gesteira.
	Programa de recuperação das demais comunidades e infraestruturas impactadas entre Fundão e Candonga	Recuperação ou reconstrução de casas, pontes e demais infraestruturas danificadas.
	Recuperação das escolas e reintegração da comunidade escolar	Garantia de acessibilidade da comunidade a escolas temporárias durante a reconstrução.
Agropecuária	Retomada das atividades aquícolas e pesqueiras	Desenvolvimento e execução de programa para apoio aos aquicultores e pescadores.
	Retomada das atividades agropecuárias	Desenvolvimento e execução de programa para o apoio aos agropecuários.
Indenizações	Levantamento e Cadastro dos Impactados	Cadastramento e avaliação dos impactos de pessoas físicas e jurídicas.
	Ressarcimento e indenização dos impactados	Execução do Programa de Indenização Mediada (PIM).
	Auxílio financeiro emergencial	Desenvolvimento de programa de auxílio financeiro emergencial à população impactada.
Gestão	Gerenciamento dos programas socioeconômicos	Monitoramento dos programas socioeconômicos através de sistemas de informação e banco de dados.
Cultura e educação	Memória histórica, cultural e artística	Recuperação de bens culturais, espaços de lazer e esporte e preservação do patrimônio histórico e cultural.
	Programa de Turismo, Cultura, Esporte e Lazer	Diagnóstico do impacto no turismo, cultura, esporte e lazer nos municípios impactados.
Animais	Programa de Assistência aos Animais	Desenvolvimento de programa para assistência aos animais extraviados e desalojados.

Fonte: Fundação Renova, 2022.

Como essas séries de iniciativas da fundação tem o potencial de modificar a situação de vulnerabilidade e o nível observado a cada período dos danos incitados aos indivíduos, este trabalho assume, como pressuposto, que projetos de mitigação (reparação) são parte do “tratamento” aplicado ao grupo de municípios atingidos. Como previram Bird e Taylor (2021), estas são reações esperadas diante deste tipo de acontecimento. Optou-se, portanto, por considerá-las como parte de um mesmo fenômeno: a destruição inicial aguda e sua subsequente reparação gradual nos anos seguintes.

Os efeitos materiais do rompimento foram muito desiguais entre as regiões atingidas. Por exemplo, o diagnóstico da Sedru (MINAS GERAIS, 2016) traz listas de efeitos identificados para cada município. O caso mais notório é que a lama de rejeitos, o efeito colateral imediato mais grave, atingiu quatro municípios (Mariana, Barra Longa, Rio Doce e Santa Cruz do Escalvado) em uma região de aproximadamente 77 quilômetros. Enquanto isso, o grupo de atingidos em nível macrorregional, segundo a classificação da Sedru, sofreu efeitos diversos na medida em que os rejeitos contaminavam o Rio Doce e seus afluentes, entre eles problemas de abastecimento de água, de esgoto e de energia elétrica. As populações ribeirinhas, que dependiam mais diretamente do uso do Rio Doce, foram consideravelmente mais afetadas do que os principais centros urbanos.

Além disto, é importante ressaltar que, para o Brasil como um todo, os anos de 2015 a 2018 formaram um período turbulento na economia nacional. Por exemplo, houve dois anos seguidos de recessão na produção interna bruta do país, o que pode ter criado um agravamento dos efeitos nestes municípios já fragilizados pelo desastre em comparação com os demais. Como se trata de um fenômeno nacional, optou-se por ignorá-lo sob a premissa de que teria afetado igualmente os demais municípios do estado. Todavia, é possível que tenha ocorrido alguma desigualdade na sensibilidade das regiões atingidas para com choques econômicos, o que poderia suscitar novos estudos futuros.

Contingente à disponibilidade pública de informação, esta avaliação se encerra com os dados de 31 de dezembro de 2018, portanto não abordando os novos mandatos das administrações estaduais e federais atuais, iniciadas em janeiro de 2019. Já as eleições municipais ocorreram em 2016, exatamente um ano após o rompimento da barragem. Da mesma forma, este estudo não abrange ano de 2020 e a pandemia de Covid-19.

O período de análise deste trabalho também não coincide com o outro rompimento de barragem em Minas Gerais, dessa vez em Brumadinho no dia 25 de janeiro de 2019, destruição semelhante que, junto com o rompimento em Mariana, passariam a ser reconhecidas como os maiores desastres socioambientais no Brasil e figurariam dentre os maiores do mundo (Rezende, 2021, p.19).

3 ÍNDICES MULTIDIMENSIONAIS

Uma vez estabelecidos os diferentes conceitos envolvidos, cabe agora descrever como foi realizado o trabalho de se mensurar os componentes familiares como fatores associados da vulnerabilidade social dos trabalhadores em condição de pobreza diante de grandes desastres.

Esta argumentação foi demonstrada a partir de um estudo de caso: o rompimento da barragem de Fundão em Mariana, MG. Este evento é um exemplo de análise do campo da demografia dos desastres. É também uma forma direta de analisar seus impactos através das características familiares como uma forma fundamental de entender os riscos que correm uma população suscetível a este tipo de perigo em função de sua composição pré-desastre.

Será esclarecido o uso de indicadores multidimensionais escaláveis de vulnerabilidade social como uma ferramenta quantitativa e útil para avaliar as relações de composição demográfica e vulnerabilidade. Neste tópico serão abordados conceitos sobre a criação de indicadores multidimensionais, suas vantagens e suas limitações. Também é apresentada uma justificativa para a opção pelo IVF, que será a base de nossa metodologia.

Em seguida, descrevem-se as bases de dados utilizadas, justificando suas opções e comentando alguns pontos de atenção necessários ao aplicá-las em análises empíricas. Por fim, as variáveis que compõem o IVF foram detalhadas, comentando como ele foi calculado face às limitações dos dados utilizados.

3.1 Procedimentos metodológicos

Este trabalho é uma análise exploratória da capacidade das famílias de mitigar riscos causados por desastres ambientais a partir de suas composições demográficas. É um estudo de caso do rompimento da barragem de Fundão, que avaliará seu impacto na vulnerabilidade social, nas dimensões de escassez de riqueza e de acesso ao trabalho, e de como este impacto varia conforme a vulnerabilidade ou a resistência das características demográficas das famílias atingidas.

Primeiro foi feita a importação e o tratamento das bases de dados, que permitiram calcular os indicadores de vulnerabilidade social, elencar os grupos de tratamento e de controle, e, por fim, observar seus resultados entre os anos de 2012 e 2018. Seguiu-se então para a definição de quais características dos municípios seriam relevantes para identificar um grupo de controle, utilizando o método *k-NN* para computar e elencar quais municípios deveriam compô-lo. Com os dois grupos

de municípios em mãos, tratou-se a base de dados do CadÚnico para mensurar anualmente seus respectivos valores para o IVF, incluindo todas as variáveis utilizadas para calculá-lo separadamente.

A avaliação foi realizada a partir da análise de diferenças em diferenças das trajetórias destes grupos. Com os resultados anuais em mãos, foi possível calcular as variações anuais e de períodos mais longos. A distância entre estas variações determinará se houve algum impacto devido àquela característica demográfica. Por exemplo: o impacto de se viver em uma família sem filhos recém-nascidos na vulnerabilidade gerada pela ausência de um ocupado em emprego formal.

No restante deste capítulo, os principais termos metodológicos para a coleta e tratamento dos dados serão explicados, enquanto a técnica de avaliação de impacto será abordada no próximo capítulo dedicado exclusivamente a ela. Por fim, teremos um capítulo demonstrando os resultados da aplicação desta metodologia para este estudo de caso.

3.2 Bases de dados

Uma vez que se trata de um desastre imprevisto, é necessário recorrer a coleta de dados secundária para as análises deste trabalho. Diante disso, não será possível elencar uma única coorte e acompanhá-la ao longo do tempo. Não há, no Brasil, pesquisas individuais ou domiciliares capazes de gerar amostras para o interior de Minas Gerais após 2010, ano do último Censo Populacional realizado até a data deste trabalho, com as informações necessárias para a construção do IVF. Foi necessário recorrer a registros administrativos que permitam calcular indicadores para a população ocupada em condição de pobreza.

3.2.1 Cadastro Único

A opção viável foi pela base desidentificada do Cadastro Único para Programas Sociais do Governo Federal (CadÚnico). Segundo o Decreto nº 11.016/2022:

"o CadÚnico é instrumento de coleta, processamento, sistematização e disseminação de informações, com a finalidade de realizar a identificação e a caracterização socioeconômica das famílias de baixa renda que residem no território nacional" (BRASIL, 2022a).

Ele é o principal registro administrativo brasileiro para os programas de assistência social, contendo uma série de informações socioeconômicas de famílias em situação de pobreza e de extrema pobreza. Um de seus objetivos estabelecido no decreto é:

"ser utilizado como repositório de dados para a realização de estudos sobre seu público, com vistas à análise de alternativas de políticas públicas para a superação da situação de vulnerabilidade econômica e social" (BRASIL, 2022a).

Em nível federal, o Ministério da Cidadania é o gestor responsável, e a Caixa Econômica Federal é o agente operador que mantém o Sistema de Cadastro Único (BASE DE DADOS, [2019]). Segundo o portal de informações da Caixa Econômica Federal ([2016]), cadastram-se apenas as famílias consideradas de baixa renda, cujo critério é um máximo de meio salário-mínimo per capita ou três salários-mínimos no total.

É possível obter publicamente a base de dados desidentificada, ou seja, uma versão ajustada para que não seja possível ferir a privacidade dos indivíduos registrados. Esta base foi anualmente atualizada, utilizando-se como referência os registros na data de 31 de dezembro de cada ano, sendo que os dados cujo acesso é imediato pelo Portal Brasileiro de Dados Abertos da Controladoria-Geral da União são os correspondentes aos anos de 2012 a 2018. Trata-se de duas bases de microdados: uma com os registros individuais e a outra com registros para a família toda (BRASIL, [2016]).

Neste estudo, fez-se uma programação em R para cruzar os dados dos indivíduos com os dados de suas respectivas famílias, realizando cálculos anuais para a população de cada grupo. Mesmo que, infelizmente, não fora possível conseguir dados mais atualizados, houve a conveniência e a oportunidade de os dados desidentificados estarem simetricamente distribuídos entre dezembro de 2012 e 2018: três anos antes e depois do rompimento da barragem do Fundão (novembro de 2015).

Além disso, foi usada a atualização de dezembro de 2014 como um ponto de partida para a avaliação do impacto em períodos distintos: em dezembro de 2015 obtiveram-se os efeitos no curto prazo, um mês após o rompimento; entre dezembro de 2016 e dezembro de 2017 são em conjunto considerados impactos de médio prazo, e os dados de 2018 representam as perspectivas para um impacto de longo prazo na condição de pobreza dos ocupados registrados no CadÚnico.

Por ser uma base de dados de registro administrativo, este cadastro está sujeito aos típicos problemas de sub-registro, qualidade das informações e da manutenção da base. Estes problemas

podem ser maiores, especialmente, por se tratar de dados sobre a parcela da população que mais carece de estrutura e acesso a serviços públicos de qualidade. Contudo, assume-se que estes problemas não sejam diferentes por regiões, ou seja, a extensão dos erros da base de dados dos municípios será a mesma ao longo dos municípios.

3.2.2 Censo 2010 e IMRS

Para a identificação do grupo de municípios de controle foram utilizadas duas outras bases, ambas calculadas em nível municipal. Uma delas foi o Censo Demográfico de 2010, onde foram calculados os indicadores mensurados por pesquisa domiciliar, considerados aqui como representativos para cada município do estado⁵. A outra fonte foram os registros administrativos compilados pela Fundação João Pinheiro para o Índice Mineiro de Responsabilidade Social (IMRS), disponibilizado para o acesso público no portal IMRS (MINAS GERAIS, [2021]). Neste segundo caso, os dados utilizados foram para o ano de 2014, por descreverem a estrutura de assistência social destes municípios um ano antes do desastre.

3.3 Indicadores multidimensionais de pobreza e de vulnerabilidade

O Índice de Desenvolvimento Humano (IDH) foi apresentado em 1990, no primeiro Relatório de Desenvolvimento Humano (RDH) produzido pela Pnud, criado a partir da colaboração entre seus principais consultadores, Mahbul ul Haq e Amartya Sen (STATON, 2007). O IDH incorpora o conceito de capacidades de Amartya Sen como parte de suas medidas de bem-estar social, adotando o argumento de Sen (1987) de que a mensuração de utilidade e bem-estar exclusivamente a partir da renda é precipitada e limitada:

“Well-being is indeed sometimes seen as reflected by the commodity-command of a person (how 'rich' he or she is), and this is one of the motivations for 'real income comparison' in terms of market command over goods and services. As an approach to interest this is,

⁵ O cálculo dos indicadores para o Censo 2010 foram gentilmente cedidos, mediante solicitação em 25 de junho de 2013, por Bernardo Alves Furtado, então Diretor Adjunto de Estudos e Políticas Regionais, Urbanas e Ambientais do Ipea, responsável por dois Textos para Discussão do Ipea sobre o Indicador de Vulnerabilidade da Família (FURTADO 2012 e 2013).

ultimately, a confusion of 'well-being' with 'being well off', and a confounding of the state of a person with the extent of his or her possessions” (SEN, 1987).

Partindo dessa perspectiva, Mahbul ul Haq estabeleceu um índice composto por três dimensões: o tradicional cálculo da renda per capita, como forma de mensurar padrões de vida; a esperança de vida ao nascer, refletindo a noção de prosperidade através de vidas longas e saudáveis, e taxas de analfabetismo e de matrícula escolar, para avaliar o acesso ao conhecimento da população. Estabeleceu-se também que as três dimensões teriam pesos iguais. O índice sofreu uma série de críticas ao longo dos anos, que motivaram melhorias técnicas e derivações do agora já tradicional IDH (STANTON, 2007).

O IDH ajudou a aumentar não só a própria discussão sobre a desenvolvimento como algo diferente de riqueza econômica, mas também a proposta de que o uso de indicadores multidimensionais simples, cujos dados podem ser obtidos globalmente, é uma excelente forma de comunicar as condições de bem-estar de homens e mulheres (STANTON, 2007).

Atualmente, há uma grande demanda pela mensuração da pobreza via indicadores multidimensionais que vão além do próprio IDH, como pode ser reconhecido na Agenda 2030 para o Desenvolvimento Sustentável e seu objetivo de erradicar a pobreza em todas as suas formas e dimensões, fazendo com que a própria Pnud se interesse por novos indicadores focados nesse objetivo (ALKIRE; JAHAN, 2019).

3.3.1 Definição

Apesar de haver ainda uma prevalência de indicadores unidimensionais, a partir da repercussão do IDH uma das principais linhas de pesquisa se voltou para a elaboração de medidas específicas para pobreza e vulnerabilidade neste espectro multidimensional. Este conceito torna muito relevante e oportuno que se calculem indicadores escalares, capazes de serem medidos desde o nível do domicílio até a comparação entre países (BARROS; CARVALHO; FRANCO, 2006; ALKIRE; FOSTER, 2011). A divergência ainda ocorre no sentido de como abordar a criação deste tipo de indicador, tanto em questões de critérios arbitrários para a escolha das dimensões da pobreza (BARROS; CARVALHO; FRANCO, 2006), quanto em relação a problemas técnicos com indicadores categóricos ou ordinais (ALKIRE; FOSTER, 2011).

3.3.2 Elaboração de índices multidimensionais

Índices multidimensionais precisam se adequar a certas características para que possuam uma capacidade de combinar elementos diversos em um mesmo denominador comum. Elementos como a relevância das dimensões, a escolha das variáveis mais adequadas, seus pesos, e suas possibilidade de escalar e desagregar por sub-regiões são o principal foco nestas discussões (BARROS; CARVALHO; FRANCO, 2006).

Chakravarty e D'Ambrosio (2013) propõem um método em que, escolhidos os indicadores, define-se uma linha quantitativa para tornar a variável binária. Com esta estrutura, uma pessoa é definida como escassa em determinado atributo se o valor do indicador original for maior ou igual à linha estabelecida, ao invés dos métodos em que a pessoa precisa ser privada de todas as características para qualificá-la como em condição de pobreza⁶.

A metodologia proposta por Chakravarty e D'Ambrosio (2013) permite elaborar indicadores com propriedades interessantes, dentre elas:

a) uma pessoa que já seja privada de algum atributo não muda seu nível de pobreza quando aumenta seu nível deste mesmo atributo. Por exemplo, uma pessoa considerada pobre devido às suas privações de saúde e segurança, mas com bons níveis de educação, não mudará sua condição de pobreza caso aumente mais seu nível de educação;

b) atributos que não sejam quantitativos dentro das dimensões e variáveis propostas é imaterial para medidas de pobreza;

c) pequenos erros de observação não geram grandes choques nos valores do índice de pobreza.

d) para qualquer formação de subgrupos da população observada, a pobreza geral se torna uma média ponderada da parcela da população dos níveis de pobreza dos subgrupos;

e) entre duas populações iguais, tudo o mais constante, aquela com linhas de privação mais altas em um atributo não deve ter uma pobreza menor devido à uma maior privação resultante de aumentos nas quantidades de subsistência;

f) a pobreza não deve aumentar caso uma pessoa mais rica migre para esta sociedade.

⁶ Para a demonstração matemática dos postulados e das propriedades destas construções de indicadores multidimensionais, veja Chakravarty e D'Ambrosio (2013).

Ambos os indicadores citados abaixo são representantes dessa linha de critérios de formulação de multidimensionalidade, mas com algumas características distintas.

3.4 Índice de Pobreza Multidimensional (IPM)

O IDH fora ainda o ponto de partida para a criação dos subsequentes indicadores de pobreza multidimensionais do Pnud, como formas de complementá-lo de uma maneira que abordasse privações e que identificasse a existência de condições de pobreza mesmo em países mais bem desenvolvidos (Pnud, 2006).

Uma antiga página do portal do Pnud na Internet (Pnud, [2012?]), descreve que o primeiro destes índices foi introduzido no Relatório de Desenvolvimento Humano de 1997 - o Índice de Pobreza Humana (IPH). O IPH buscava agregar características diferentes de privação e medir a extensão da pobreza em uma comunidade. Derivando suas dimensões para refletir aquelas que compõem o IDH, ele focava em variáveis de sobrevivência, de conhecimento e de padrões de vida decentes. Era dividido em dois métodos, o IPH-1 para países em desenvolvimento, e o IPH-2 para os países membros da OCDE, em uma tentativa de visualizar melhor condições de pobreza internas em países que, em média, já possuíam um alto nível de riqueza⁷.

Segundo Barros, Carvalho e Franco (2006), estes índices possuíam algumas dificuldades em suas possibilidades de escala. Eles não eram capazes de alcançar uma desagregação familiar, limitando-se a gerar níveis médios regionais:

“Podemos calcular os IPHs de um país, de uma cidade ou mesmo de um bairro, mas não podemos calcular os IPHs de uma família ou dos negros ou das mulheres em uma dada comunidade. Isso ocorre porque, no cálculo dos IPHs, primeiro se agregam as informações de cada uma das dimensões da pobreza para as famílias de uma dada área.” (BARROS; CARVALHO; FRANCO, 2006)

Eles sofreram muitas críticas também pelo uso de poucas variáveis, fragilizando a aproximação entre resultados quantitativos dos indicadores e os objetivos declarados de cada dimensão (BARROS; CARVALHO; FRANCO, 2006).

⁷ O IPH-2 também possuía uma dimensão para mensurar exclusão social, baseada em taxas de desemprego prolongado [PNUD, [2012?]]

A partir de 2010, o Pnud criou, em colaboração com a Oxford Poverty e a Human Development Initiative (OPHI), um novo indicador, o Índice de Pobreza Multidimensional (IPM), qualificando melhor a metodologia de um cálculo para identificar de forma comparativa condições de pobreza agudas e experiências de privações múltiplas de forma escalar (ALKIRE; JAHAN; 2018).

Desta vez, ainda se espelhando nas três principais dimensões do IDH (saúde, educação, e padrão de vida), o IPM passou a contar com 10 indicadores, tornando seus dados mais robustos que os anteriores (ALKIRE; JAHAN, 2018; Pnud, 2019).

O IPM utilizar a metodologia inspirada naquela elaborada por Alkire e Foster (2011), o que trouxe uma robustez técnica significativa e, principalmente, a possibilidade de escalar medidas de pobreza em nível micro, pois parte de variáveis mensuradas no nível da família e do indivíduo, ao invés de dados agregados como ocorriam com os IPH (Ophi; Pnud, 2019).

Tabela 3-1 – Dimensões e indicadores que compõem o IPM

IPM	Saúde	Nutrição
		Mortalidade infantil
	Educação	Anos de escolaridade
		Adesão escolar
	Padrão de vida	Combustível para cozinhar
		Saneamento
		Água potável
		Eletricidade
		Habitação
		Ativos

Fonte: Oxford Poverty e Human Development Initiative (2018).

3.4.1 Método Alkire-Foster

O método Alkire-Foster (AF) de mensuração de pobreza multidimensional foi desenvolvido com bases nas medidas de pobreza Foster-Greer-Thorbecke (FGT), e calculando uma relação de situações de privação sofridas por uma pessoa em um dado, através de dimensões e indicadores originários que subsidiam seu cálculo (OPHI, 2015).

Uma de suas vantagens é a capacidade de analisar tanto a proporção da população em condições de pobreza quanto o nível de intensidade sofrido. Ele também é capaz de se decompor

em subgrupos dentro da população avaliada, mesmo que não sejam grupos regionais, mas a partir de outras características, como raça ou idade.

Uma pessoa é considerada em estado de pobreza multidimensional quando a soma de suas privações se tornar maior ou igual uma quantidade (ou proporção) de indicadores que se situam abaixo linha de pobreza estabelecida para cada indicador. Cada variável de privação possui uma característica binária: para cada variável analisada, é preciso estabelecer um limite de valor, e então a nova variável será simplesmente dividida entre privação (indicador pior do que o limite) e não privação (indicador melhor que o limite).

Segundo o relatório Global Multidimensional Poverty Index 2019 (PNUD, 2019), para o caso do IPM há o produto entre a incidência (H) e a intensidade da pobreza multidimensional (A). A pessoa será definida em condição de pobreza multidimensional um terço ou mais de seus indicadores identificaram privações, e severamente pobre caso seja metade ou mais ($H \times A$).

Este se trata de uma das possibilidades do AF, definida como *Adjusted headcount ratio* (M_0), sendo que os autores apresentam também possibilidades de se medir como *Adjusted Poverty Gap* (M_1), que inclui um produto que calcula a diferença entre a condição de pobreza observada com a estabelecida na linha de corte ($H \times A \times G$). Por fim, a medida *Adjusted Squared Poverty Gap* (M_2) inclui uma sensibilidade para os diferentes níveis de privação dentre a população pobre ($H \times A \times S$), ao invés de simplesmente comparar as médias, permitindo uma estimativa da própria desigualdade entre as privações sofridas na população (ALKIRE; FOSTER, 2011)⁸.

3.5 Índice de Vulnerabilidade das Famílias (IVF)

Uma segunda abordagem para a elaboração de indicadores multidimensionais foi utilizada por Barros, Carvalho e Franco (2006). Suas características são inspiradas em trabalhos anteriores de Chakravarty, e estão alinhadas com os postulados que Chakravarty e D'ambrosio (2013) descrevem.

⁸ Apenas o método M_0 permite o uso de variáveis não cardinais. Para a demonstração matemática dos cálculos de cada método, veja Alkire e Foster (2011).

Em um primeiro momento, este indicador recebera o nome de Índice de Desenvolvimento da Família (IDF), mas a partir de 2012 Furtado (2012; 2013) o definiria como Índice de Vulnerabilidade das Famílias (IVF)⁹.

O IVF foi elaborado baseando-se nas características das duas principais pesquisas domiciliares brasileiras realizadas pelo IBGE, o Censo brasileiro e a Pesquisa Nacional por Amostra de Domicílios (Pnad), o que permite uma análise histórica para o Brasil e para suas diferentes regiões geográficas e populações, com características diversas e desiguais (BARROS; CARVALHO; FRANCO, 2006). Assim como os demais indicadores de pobreza multidimensionais citados, o IVF tem como estratégia principal ponderar variáveis de indicadores sociais de uso mais comuns.

Além disso, buscando mitigar as críticas aplicadas ao IPH, os autores utilizaram um número maior de variáveis para formar o indicador sintético, mas que ainda assim preservasse sua natureza escalar. A principal crítica é pela origem do cálculo dos indicadores. Barros, Carvalho e Franco (2006) argumentam que um problema dos IPH era que sua mensuração partia de indicadores agregados por região, e não por agente (indivíduo ou família, por exemplo). Isso limitaria sobremaneira as capacidades de desagregação do índice, especialmente para subgrupos não baseados em regiões geográfica, como distribuição por idade ou sexo.

O IVF tem como uma de suas características a linearidade, e não necessita de um valor agregado de “condição de pobreza”, podendo ser utilizado primeiro como uma medida agregada de bem-estar:

“O princípio dessa abordagem alternativa seria, em vez de buscar obter um indicador agregado da pobreza de cada indivíduo – o que requereria, entre outras coisas, linhas de pobreza para cada dimensão –, obter um agregado de bem-estar e a partir daí obter a distribuição dos agentes segundo seu nível de bem-estar. Nesse caso, o agregado de pobreza deveria ser definido como uma característica da cauda dessa distribuição, por exemplo, a proporção dos agentes com nível de bem-estar abaixo de um nível mínimo ou o equivalente ao hiato de pobreza.” (BARROS; CARVALHO; FRANCO, 2006)

⁹ Eventualmente o Ipea focaria seus esforços em um novo indicador multidimensional, o Índice de Vulnerabilidade Social, que não será abordado neste trabalho. Ver Costa e Marguti (2015).

3.5.1 Vulnerabilidade das famílias

Um ponto crucial na formação do IVF foi a opção pelo cálculo a partir de níveis de vulnerabilidade social, ao invés de especificamente observar a pobreza das famílias. Desta forma, o índice procura estimar as dificuldades de resposta das famílias para qualquer sorte de adversidades socioambientais, sejam características específicas de cada região, ou sejam suas instituições, capital social, condições econômicas dos habitantes do local (FURTADO, 2013).

Por exemplo, considere duas famílias A e B, idênticas em todos os aspectos relevantes para vulnerabilidade: cada uma com um casal de adultos; mesmo número de filhos; mesmo emprego; residências e demais características em comum. **A única diferença entre elas é que** a família A possui um membro a mais em seu domicílio, como um idoso que precise de cuidados especiais. Neste caso, a família A é considerada mais vulnerável que a família B, ainda que nenhuma das duas sejam necessariamente pobres em nível multidimensional. Diante de um cenário súbito adverso, como a perda de emprego do responsável, a família B teria melhores condições de buscar um novo emprego, porque a família A ainda precisa dedicar tempo e esforço para a pessoa com necessidades especiais enquanto procura as mesmas oportunidades de voltar a trabalhar.

Furtado (2013) também define que o sentido de vulnerabilidade de famílias adotado na construção do IVF baseou-se nos conceitos de risco e vulnerabilidade da defesa civil, bem como capacidades de resposta a eventos adversos inesperados na população.

3.5.2 Construção do índice

Barros, Carvalho e Franco (2006) deram ao índice seis dimensões, subdivididas em 26 componentes e estes, por sua vez, em 48 variáveis. A formatação das variáveis segue uma estrutura binária: a partir de uma ou mais variáveis originais da pesquisa, estabelece-se um critério como uma linha de corte, em que a pessoa é considerada vulnerável quando cumpre os critérios definidos. Assim, uma nova variável emerge com a condição de vulnerável sendo igual a 1, e a condição de não-vulnerável (resistente) sendo igual a 0.

É fundamental destacar a polaridade menor-melhor do indicador. Ou seja, os critérios estabelecidos buscam sempre identificar a vulnerabilidade, e não a resistência. Quanto maior o valor agregado do indicador, mais vulnerável é a condição do agente (FURTADO, 2013).

Isto é importante ressaltar porque assume-se uma postura diferente de indicadores como o IDH ou o IPM, que mantêm sua polaridade maior-melhor. Furtado (2012; 2013) utiliza como resultados uma variação decimal entre 0 e 1, assim como o IDH, mas um IFV igual a 0 significa absolutamente resiliente, enquanto 1 significa absolutamente vulnerável (BARROS; CARVALHO; FRANCO, 2006; FURTADO, 2012; 2013)¹⁰.

Outro aspecto relevante é que, apesar das 6 dimensões possuírem pesos iguais, elas são formadas por componente e variáveis que podem, por sua vez, receber pesos num formato hierárquico de médias. Além disso, apesar do indicador permitir resultados até para cada indivíduo, seus critérios são todos baseados em variáveis para todo o domicílio (BARROS; CARVALHO; FRANCO, 2006).

Buscando clarificar estes últimos pontos de atenção, abaixo será realizada uma descrição mais detalhada do método de cálculo do indicador, bem como suas dimensões e variáveis, e a necessidade de adaptação para o cálculo pelo CadÚnico ao invés de Censo ou Pnad.

3.6 Dimensões e variáveis do IVF

Barros, Carvalho e Franco (2006) e Alkire e Foster (2011) argumentam que ainda não há um consenso sobre quais deveriam ser as dimensões e os pesos para um indicador multidimensional de pobreza ou vulnerabilidade:

“However, as a general methodology for identifying the poor, the dual cutoff approach has not been explicitly formulated in the literature, nor have its implications for multidimensional poverty measures – or their axioms – been explored.

[...]

Yet in other settings there may be good arguments for using general weights. Indeed, the choice of dimensional weights may be seen as a value judgement which should be open to public debate and scrutiny”. (ALKIRE; FOSTER, 2011)

“Seguramente, o primeiro passo na construção de uma medida de pobreza consiste sempre em definir quais dimensões são mais relevantes e quais devem ser os indicadores utilizados para representar cada uma. De todas as etapas do processo de construção de um indicador de pobreza multidimensional, essa é uma das em que a literatura menos avançou. Apesar de importantes contribuições como a de Sen (1992) e a de Dasgupta (1993), ainda

¹⁰ Na primeira versão do indicador utilizada por Barros, Carvalho e Franco (2006), os autores definiram o indicador de modo que o resultado final ficasse entre 0 a 100.

estamos longe de saber normativamente quais são as dimensões mais relevantes da pobreza e também distantes de ter definido os melhores indicadores para representar cada dimensão.” (BARROS; CARVALHO; FRANCO, 2006; FURTADO, 2012; 2013)

Foi nesse sentido que Barros, Carvalho e Franco (2006) optaram por criar as dimensões focando no pragmatismo, ao considerar a oportunidade de integrar o indicador facilmente à Pnad. Além disso, argumentam que seria difícil encontrar um cálculo de peso que gerasse uma hierarquia entre as dimensões, optando por considerar todas com o mesmo peso para o índice.

Em última instância, estas dimensões representam uma perspectiva das capacidades como meio para se satisfazerem as necessidades básicas, na medida em que a vulnerabilidade representa o risco de faltar o acesso aos meios de se resistir a adversidades. Barros, Carvalho e Franco (2006) lamentam que, das condições de pobreza tradicionalmente consideradas, o tema da saúde não pode ser incluído de forma satisfatória com base nos dados da Pnad.

As seis dimensões podem ser definidas como¹¹:

- i) vulnerabilidade demográfica (V);
- ii) acesso ao trabalho (T);
- iii) escassez de recursos (R);
- iv) acesso ao conhecimento (C);
- v) desenvolvimento infantojuvenil (D); e
- vi) condições habitacionais (H)¹².

Conforme explicitado anteriormente, cada variável será binária, com valor 1 para vulnerável e 0 para resiliente ou não-vulnerável. No entanto, alguns atributos aparecem em mais de uma variável. Por exemplo, enquanto a variável V1 identifica a presença de filho nascido vivo

¹¹ O nome utilizado pelos autores para a dimensão de vulnerabilidade demográfica foi “vulnerabilidade social”. No entanto, para este trabalho, esse nome soaria incompleto e confuso. Todos as variáveis que a compõem são eminentemente demográficas, tratando de nascidos vivos, dependentes por faixas etárias, idade ativa, nupcialidade e maternidade. Isto posto, denominá-la como vulnerabilidade demográfica permitirá um melhor entendimento de que esta será nosso principal foco de discussão do IVF como um todo. Por fim, o termo “vulnerabilidade social” foi amplamente discutido neste trabalho, e seu uso aqui exigiria mais critérios do que os originais propostos por Barros, Carvalho e Fraco (2006).

¹² A ordem de apresentação das dimensões foi alterada em comparação com a original. Barros, Carvalho e Franco (2006) ordenaram a apresentação das dimensões por um critério de associação de quais são mais ligadas a meios de se tornar resiliente e quais se destacam como fins. Na prática, isso não afeta em nada o cálculo do índice e destas dimensões. A ordem definida neste trabalho serve apenas uma questão de organização das informações baseada no foco do estudo em demografia e ocupação em condições de pobreza.

no último ano; a V2 busca identificar a presença de filhos nascidos vivos nos últimos dois anos. Assim, apesar de não haver peso entre as seis dimensões, as variáveis utilizadas para gerar seus valores passam por um efeito de “indicadores de cascata”:

“O princípio utilizado torna-se idêntico ao do hiato de pobreza se considerarmos que um adulto tem um indicador de falta de vulnerabilidade superior a 3, que o dos jovens é 2, dos adolescentes 1 e das crianças 0. Nesse caso, se a linha de vulnerabilidade é 3, então a distância de cada jovem à linha de vulnerabilidade é 1, a de cada adolescente é 2 e de cada criança 3”. (BARROS; CARVALHO; FRANCO, 2006)

As definições de cada dimensão serão abordadas abaixo, discutindo as variáveis utilizadas originalmente, mais as eventuais adaptações causadas pela limitação dos dados do CadÚnico.

3.6.1 Vulnerabilidade demográfica (V)

Esta dimensão trata da composição da família, visando avaliar as condições de vulnerabilidade causadas pela dinâmica demográfica entre seus membros. Temos em seus aspectos basicamente as análises de fecundidade, composição etária, dependência, nupcialidade, maternidade e morbidade.

A fecundidade considera vulnerável a presença de mulher que teve filho nascido vivo no último ano (V1), ou nos últimos dois anos (C2).

A necessidade de cuidado e dedicação aos dependentes compõem as variáveis V3 (criança), V4 (criança ou adolescente), V5 (criança, adolescente ou jovem) e V6 (idosos com 65 anos ou mais). Há ainda um cálculo de taxa de dependência “intrafamiliar” (V8), que assume como vulnerável uma família cuja minoria nos seus membros esteja em idade ativa.

Um ponto interessante é a análise de vulnerabilidade pela ausência de um cônjuge no domicílio (V7), focando justamente em uma das composições familiares que Polizzi, Struffolino e Van Winkle (2020) identificaram como determinante de entrada na condição de pobreza.

Um último aspecto proposto por Barros, Carvalho e Franco (2006) é a ausência especificamente da mãe, através das variáveis V9 (criança cuja mãe tenha morrido) e V10 (criança que viva em um domicílio sem a presença da mãe).

Houve duas adaptações na versão de Furtado (2013). Ele elimina a V9, considerando, portanto, apenas se a mãe é ausente, independente do motivo; além disso, inclui uma nova variável

para este nome V9, que define como vulnerável a família que possua como membro alguma pessoa com dificuldade grande ou impossibilidade de caminhar, enxergar, ouvir ou com deficiência mental permanente.

Dadas as limitações do CadÚnico, optou-se neste trabalho por manter a versão de Furtado (2013), incluindo a condição de deficiência (como é definida na base desidentificada do CadÚnico), e excluindo a variável V10.

Esta dimensão é fundamental para esta avaliação, pois será a partir dela que poderemos avaliar os efeitos da composição familiar na vulnerabilidade dos ocupados em condição de pobreza.

Tabela 3-2 – Variáveis componentes da dimensão Vulnerabilidade demográfica, conforme adaptadas para a base desidentificada do CadÚnico

Indicador	Critério para condição de vulnerabilidade
V1. Alguma mulher teve filho nascido vivo no último ano	Presença de filho de idade 0 em 31/dez do ano de referência.
V2. Alguma mulher teve filho nascido vivo nos últimos dois anos.	Presença de filho de idade 0 ou 1 em 31/dez do ano de referência.
V3. Presença de criança.	Presença de pessoa com até 12 anos de idade
V4. Presença de criança ou adolescente.	Presença de pessoa com até 17 anos de idade
V5. Presença de criança, adolescente ou jovem.	Presença de pessoa com até 21 anos de idade
V6. Presença de idoso	Presença de pessoa com 65 ou mais anos de idade
V7. Ausência de cônjuge.	Ausência de pessoa definida como cônjuge da pessoa de referência
V8. Menos da metade dos membros encontram-se em idade ativa	Pessoas com 10 anos ou acima são minoria no domicílio
V9. Presença de pessoas com dificuldade grande ou impossibilidade de caminhar, enxergar, ouvir ou com deficiência mental permanente.	Presença de alguém declarado como possuindo deficiência
V10. Presença criança no domicílio que não viva com a mãe	Considera a presença de criança (como V3), a ausência de cônjuge (V7), e a ausência de adultos mulheres

Fonte: Adaptado de Furtado (2013)

3.6.2 Acesso ao trabalho (T)

O acesso ao trabalho possui três identificadores de condição de vulnerabilidade. As duas primeiras variáveis avaliam a disponibilidade de trabalho oferecida pelo mercado e atendida pela família. T1 avalia a desocupação, sendo vulnerável o domicílio com mais da metade das pessoas

em idade ativa desocupadas, enquanto T2 depende de que ninguém esteja a mais de seis meses no emprego atual.

Outras duas variáveis estão mais relacionadas com a qualidade do posto do trabalho: T3 analisa a ausência de ocupados formalizados e T4 considera a ausência pessoas em setores econômicos não-agrícolas como uma vulnerabilidade.

Por fim, T5 e T6 consideram a remuneração dos ocupados, tratando como vulneráveis os domicílios que não possuam nenhum indivíduo recebendo mais de um ou dois salários-mínimos, respectivamente. Esta é uma análise diferente daquelas previstas na dimensão de escassez de renda, discutida abaixo. Estas variáveis são focadas na renda individual: como se trata da condição de salário do empregado, considerando que qualquer ocupação que não remunere ao menos um salário-mínimo se trata de um trabalho informal e, portanto, sem acesso às políticas de proteção ao trabalhador, como o seguro-desemprego. Nesse sentido, trata-se de variáveis mais focadas na qualidade do trabalho do que na renda do domicílio como um todo, contemplada em uma dimensão própria.

A base desidentificada do CadÚnico não permite o cálculo da variável T2, uma vez que não divulga quanto tempo a pessoa já está empregada. Optou-se por descartá-la do indicador.

Tabela 3-3 – Variáveis componentes da dimensão Acesso ao trabalho, conforme adaptadas para a base desidentificada do CadÚnico

Indicador	Critério para condição de vulnerabilidade
T1. Menos da metade dos membros em idade ativa encontram-se ocupados	Pessoas em idade ativada ocupadas são minoria no domicílio
T3. Ausência de ocupado no setor formal	Todas as pessoas ocupadas na semana de referência estão em uma dessas condições: empregado sem carteira assinada; conta-própria; temporário em área rural; não remunerado
T4. Ausência de ocupado em atividade não-agrícola	Todas as pessoas ocupadas na semana de referência em atividades extrativistas
T5. Ausência de ocupado com rendimento superior a um salário-mínimo	Valor de remuneração do emprego menos que um salário-mínimo
T6. Ausência de ocupado com rendimento superior a dois salários-mínimos	Valor de remuneração do emprego menos que dois salários-mínimos

Fonte: Adaptado de Furtado (2013)

3.6.3 Escassez de recursos (R)

A escassez de recursos mostra efeitos importantes, pois a condição de pobreza é de requisito para o registro no Cadastro Único, a partir da renda per capita ou total abaixo das linhas definidas pelo governo federal. Nesse sentido, é de se esperar que se trate dos maiores níveis de vulnerabilidade em um indicador multidimensional.

Há apenas três indicadores, sendo dois deles fortemente correlacionados: extrema pobreza (R1) e pobreza (R2). Eles implicam na condição da renda per capita do domicílio, considerando todas as suas fontes de renda declaradas. A condição de pobreza está definida como uma renda per capita inferior a meio salário-mínimo, conforme adotado no ano analisado e estimando seus valores reais a partir do Índice de Preços ao Consumidor Amplo (IPCA). Já a extrema pobreza ocorre em domicílios com renda per capita de 1/4 do salário-mínimo.

O terceiro indicador busca identificar a qualidade da fonte de renda do domicílio como um todo. Calculou-se a renda total por todas as fontes, para depois identificar qual a participação da renda advinda de trabalho, em contrapartida a outras fontes como transferências sociais ou doação. A família será considerada vulnerável caso a minoria de sua renda seja através de trabalho dos seus membros (R3).

O fato de o CadÚnico se limitar à população em condição de pobreza, ainda que em um critério que difere daquele para o Bolsa-Família, inviabiliza uma análise da entrada e saída das famílias nessa condição, como calculado por Polizzi, Struffolino e Van Winkle (2020). Além de trazer o necessário alerta de que, qualquer que seja a avaliação e seu resultado neste trabalho, ele não permitirá criar inferências para toda a população dos municípios atingidos, apenas para o universo de pessoas que já possuem um perfil de maior vulnerabilidade.

Tabela 3-4 – Variáveis componentes da dimensão Escassez de recursos, conforme adaptadas para a base desidentificada do CadÚnico

Indicador	Critério para condição de vulnerabilidade
R1. Renda familiar per capita inferior à linha de extrema pobreza	Renda per capita de até 1/4 do salário-mínimo vigente
R2. Renda familiar per capita inferior à linha de pobreza	Renda per capita até 1/2 do salário-mínimo vigente
R3. Maior parte da renda familiar advém de transferências	Valor de remuneração do emprego menor que a soma das demais fontes de remuneração

Fonte: Adaptado de Furtado (2013).

3.6.4 Acesso ao conhecimento (C)

A dimensão do acesso ao conhecimento analisa as condições de escolaridade dos adultos do domicílio: a escolaridade dos mais jovens será analisada na dimensão de desenvolvimento infantojuvenil. Ou seja, esta trata do grau de capital humano que já afeta as condições de pobreza da família, enquanto a próxima foca prioritariamente nas possibilidades de investimento de capital humano e dedicação ao cuidado dos dependentes jovens. Ressalta-se que estas variáveis consideram como adultos as pessoas com 17 anos ou mais.

As seis variáveis desta dimensão podem ser basicamente divididas em três partes. A primeira analisa a presença de familiares com 17 anos ou mais, analfabetos ou analfabetos funcionais no domicílio (C1 e C2, respectivamente), considerando as duas formas de analfabetismo como presença de vulnerabilidade. Já a segunda trata da importância de se ter adultos com graus de escolaridade fundamental (C3) e médio (C4) completos, ou com nível superior (C5), ainda que incompleto.

A última parte, qualificação profissional, muda a avaliação exclusivamente para os ocupados no domicílio, com uma única variável: a ausência de trabalhador com qualificação média ou alta (C6). Esta variável precisou ser calculada de forma diferente da proposta original. A variável utilizada no Censo parte da classificação da ocupação, segundo os grandes grupos da Classificação Brasileira de Ocupações (CBO) de 2002 (Ipea, 2010). Ela então define trabalhador de qualificação média aquele que possua um ofício classificado como "dirigentes em geral, profissionais das ciências e das artes e técnicos de nível médio (excluídos oficiais forças armadas)" (Furtado, 2013). Esta informação não está disponível na base desidentificada do Cadastro Único, portanto, optou-se por uma aproximação baseada na qualificação do próprio ocupado, não de sua função no trabalho. Esta estratégia limita a capacidade de analisar a qualidade do emprego da pessoa, mas permite estimar quais são suas habilidades para lidar com o seu trabalho.

Tabela 3-5 – Variáveis componentes da dimensão Acesso ao conhecimento, conforme adaptadas para a base desidentificada do CadÚnico

Indicador	Critério para condição de vulnerabilidade
C1. Presença de adulto analfabeto	Pessoa não sabe ler ou escrever
C2. Presença de adulto analfabeto funcional	Pessoa maior de 17 anos e ensino fundamental incompleto
C3. Ausência de adulto com fundamental completo	Pessoa maior de 17 anos e ensino fundamental completo ou maior
C4. Ausência de adulto com secundário completo	Pessoa maior de 17 anos e ensino médio completo ou maior
C5. Ausência de adulto com alguma educação superior	Pessoa maior de 17 anos e ensino superior incompleto ou maior
C6. Ausência de trabalhador com qualificação média ou alta	Pessoa maior de 17 anos ocupado e com ensino médio completo ou maior

Fonte: Adaptado de Furtado (2013)

3.6.5 Desenvolvimento infantojuvenil (D)

A dimensão de desenvolvimento infantojuvenil, como citado acima, concentra as variáveis de educação reservadas às crianças e aos jovens menores de 17, além da presença de trabalho infantil no domicílio.

Nesse sentido, as duas primeiras variáveis (D1 e D2) tratam da presença de menores de 14 e de 16 anos ocupados, respectivamente. As três variáveis seguintes identificam a presença de crianças ou jovens fora da escola, até 6 anos (D3), de 7 a 14 (D4) e de 7 a 17 anos (D5). A variável D6 faz análise parecida, mas identifica a defasagem escolar ao invés de uma evasão ou de abandono. Por fim, as duas últimas variáveis, D7 e D8, procuram identificar a presença de analfabetos de 10 a 14 anos e de 15 a 17 anos, respectivamente.

Por fim, D9, D10 e D11 consideram os casos em que houve a morte de filhos na família: a presença de ao menos uma mãe (D9) ou mais de uma mãe (D10), que tenha algum filho que já tenha morrido, enquanto D11 avalia os casos de filhos nascidos mortos.

Como não há dados sobre mortalidade na base desidentificada do Cadastro Único, não foi possível analisar as variáveis: presença de ao menos uma mãe que tenha algum filho que já tenha morrido (D9); presença de mais de uma mãe que tenha algum filho que já tenha morrido (D10); e presença de mãe que já teve algum filho nascido morto (D11). Estas análises precisaram ser ignoradas por não haver sequer alguma forma de aproximação destas informações.

Tabela 3-6 – Variáveis componentes da dimensão Desenvolvimento infantojuvenil, conforme adaptadas para a base desidentificada do CadÚnico

Indicador	Critério para condição de vulnerabilidade
D1. Presença de ao menos uma criança com menos de 14 anos trabalhando	Presença de ocupado na semana com 14 anos ou menos
D2. Presença de ao menos uma criança com menos de 16 anos trabalhando	Presença de ocupado na semana com 16 anos ou menos
D3. Presença de ao menos uma criança de 0-6 anos fora da escola	Pessoa com até 6 anos de idade nunca frequentou ou não frequenta mais a escola
D4. Presença de ao menos uma criança de 7-14 anos fora da escola	Pessoa de 7 a 14 anos de idade nunca frequentou ou não frequenta mais a escola
D5. Presença de ao menos uma criança de 7-17 anos fora da escola	Pessoa de 7 a 17 anos de idade nunca frequentou ou não frequenta mais a escola
D6. Presença de ao menos uma criança de até 14 anos com mais de 2 anos de atraso	Pessoa de até 14 anos estando 2 anos ou mais de idade acima da idade esperada para o seu nível de escolaridade
D7. Presença de ao menos um adolescente de 10 a 14 anos analfabeto	Pessoa entre 10 a 14 anos é analfabeta (C1) ou analfabeta funcional (C2).
D8. Presença de ao menos um jovem de 15 a 17 anos analfabeto	Pessoa entre 15 a 17 anos é analfabeta (C1) ou analfabeta funcional (C2).

Fonte: Adaptado de Furtado (2013)

3.6.6 Condições de habitação (H)

Na dimensão de habitação, as variáveis tratam, por um lado, da infraestrutura que podem ter uma relação macro com sua região no município e, por outro, questões específicas do lar. A análise da infraestrutura compreende a disponibilidade adequada de tratamento de água (H5) e de esgoto (H6), bem como de coleta de lixo (H7) e de energia elétrica (H8). As questões mais específicas tratam de condições como o domicílio ser coletivo (H1) ser improvisado (H2), ou ainda possuir uma densidade de 2 ou mais moradores por dormitório (H8).

No Cadastro Único, quando comparado com os dados disponíveis no Censo, esta dimensão sofreu alteração e perda de variáveis. A variável H1, domicílio coletivo, tenta substituir duas variáveis originais: se o domicílio é não é próprio ("pago", "pagando"), ou ainda se não é nem próprio nem cedido (alugado ou em outra condição). Havia também variáveis que identificavam bens possuídos: geladeira (H9 a H12), televisão (H10 a H12), rádio (H10 a H12), telefone (H11 a H12) ou computador (H12). Não havia também a disponibilidade de informação sobre comutação intermunicipal (H13, Presença de membro que "trabalha ou estuda em município diferente da residência"). Por fim, a variável sobre o domicílio ser improvisado era originalmente H14 ao invés de H2.

Tabela 3-7 – Variáveis componentes da dimensão Condições de habitação, conforme adaptadas para a base desidentificada do CadÚnico

Indicador	Critério para condição de vulnerabilidade
H2. Domicílio é coletivo	Espécie de domicílio coletivo
H3. Densidade de dois ou mais moradores por dormitório	Quantidade de pessoas maior ou igual a 2 por cômodos servindo como dormitório do domicílio
H5. Acesso inadequado à água	Se o domicílio não tem água encanada
H6. Esgotamento sanitário inadequado	Forma de escoamento sanitário não se dá por fossa séptica ou rede coletora de esgoto ou pluvial
H7. Lixo não é coletado	Lixo não é coletado direta ou indiretamente
H8. Sem acesso à eletricidade	Tipo de iluminação não é elétrica, seja com ou sem medidor próprio ou comunitário
H14. Domicílio improvisado	Espécie de domicílio particular improvisado

Elaboração Própria. Adaptado de Furtado (2013)

3.7 Agregação das variáveis e dimensões

Para o IVF, optou-se pela agregação linear em pesos únicos por dimensão. Contudo, sua construção por variáveis se propôs a computar as variáveis de vulnerabilidade em um formato de índice em cascata, a fim de permitir que condições de vulnerabilidade muito similares (ex. como faixas etárias ou salariais) pudessem se mostrar como mais ou menos agravantes (ex. um filho de 1 ano provoca uma vulnerabilidade maior que um filho de 12 anos).

Mantém-se aqui o racional proposto por Barros, Carvalho e Franco (2006) para o cálculo do IVF¹³. Considerando:

S o valor do Índice de Vulnerabilidade da Família de um dado indivíduo;

B_k o valor de uma dimensão k , em uma série $\{B_k: k = 1, \dots, m\}$;

w_k o peso dado ao indicador B_k ;

w_k igual a 1 para todas as dimensões;

B_{ijk} o i -ésimo indicador básico do j -ésimo componente da k -ésima dimensão;

m_k o número de variáveis da k -ésima dimensão;

n_{jk} o número de componentes da j -ésimo variável da k -ésima dimensão;

6 o número de dimensões;

1 o peso;

Decorre que:

¹³ Trata-se de um resumo; para ver a demonstração completa da expressão, ver Barros, Carvalho e Franco (2006)

$$S = \sum_{k=1}^6 \sum_{j=1}^{m_k} \sum_{i=1}^{n_{jk}} \frac{B_{ijk}}{6 \cdot m_k \cdot n_{jk}}$$

Considerando suas propriedades, e o peso 1, tem-se que:

$$S = \frac{1}{6} \sum_{k=1}^6 S_k$$

Para agregar a vulnerabilidade por família e grandes grupos, pode-se denotar:

f = família

$S(f)$ = vulnerabilidade da família f

P = vulnerabilidade média das famílias do subgrupo;

Podemos usar a expressão para o grau de vulnerabilidade da sociedade como

$$P = \frac{1}{6} \sum_{k=1}^6 \left(\sum_f \sum_{j=1}^{m_k} \sum_{i=1}^{n_{jk}} \frac{V_{ijk}(f)}{N \cdot m_k \cdot n_{jk}} \right) = \frac{1}{6} \sum_{k=1}^6 P_k$$

Onde V_k equivale à vulnerabilidade da sociedade referente à dimensão k nas 40 variáveis, sendo todas binárias. Além disso, é possível estabelecer cada variável de vulnerabilidade que possua mais de um componente como:

$$P_{jk} = \sum_f \sum_{i=1}^{n_{jk}} \frac{B_{ijk}(f)}{N \cdot n_{jk}}$$

3.7.1 Expressões para cada dimensão

Cada dimensão foi formulada considerando a seguinte distribuição das variáveis em seus componentes:

a) Vulnerabilidade demográfica:

$$V = \left[\frac{(V_1 + V_2)}{2} + \frac{(V_3 + V_4 + V_5)}{3} + V_6 + \frac{V_7 + V_8}{2} + V_9 + V_{10} \right] / 6$$

b) Acesso ao trabalho:

$$T = \left[T_1 + \frac{(T_3 + T_4)}{2} + \frac{T_5 + T_6}{2} \right] / 3$$

c) Escassez de recursos:

$$R = (R_1 + R_2 + R_3) / 3$$

d) Acesso ao conhecimento:

$$C = \left[\frac{(C_1 + C_2)}{2} + \frac{(C_3 + C_4 + C_5)}{3} + C_6 \right] / 3$$

e) Desenvolvimento infantojuvenil:

$$D = \left[\frac{(D_1 + D_2)}{2} + \frac{(D_3 + D_4 + D_5)}{3} + \frac{(D_6 + D_7 + D_8)}{3} \right] / 3$$

f) Condições habitacionais:

$$H = (H_1 + H_2 + H_3 + H_4 + H_5 + H_6 + H_7 + H_8) / 8$$

4 AVALIAÇÃO DE IMPACTO POTENCIAL ROMPIMENTO DA BARRAGEM NA VULNERABILIDADE SOCIAL

Quando se quer melhorar o alcance de objetivos através da execução de políticas públicas, é necessário recorrer a métodos de monitoramento e avaliação, aqui definidos como as técnicas de mensuração das qualidades e dos defeitos da própria política: de uma forma análoga a como outras ciências estudam os efeitos de suas aplicações; por exemplo, como as ciências bioquímicas lida com a produção de remédios para o tratamento de doenças.

Técnicas de monitoramento são aquelas utilizadas ao longo da execução da política, enquanto técnicas de avaliação são aquelas que estudam a política antes ou depois de cada um de seus principais marcos (planejamento, seleção, execução de cada fase, conclusão e impactos). Como são dois conceitos tão próximos e por vezes confundidos um com o outro, analistas de políticas tendem a definir o conjunto das técnicas de metodologias de M&A (monitoramento e avaliação; UNICEF, 1995).

Neste capítulo, descreve-se em detalhes o conceito e marco teórico de um tipo de M&A específico, conhecido como avaliação de impacto. Depois de uma visão geral, faz-se um detalhamento do método de avaliação de impacto utilizado neste trabalho, as diferenças em diferenças.

Em seguida, tem-se a descrição de como foi estruturada a avaliação de impacto desta dissertação, em especial acerca da estratégia de seleção dos municípios elencados para esta análise.

4.1 Definindo avaliação de impacto

Uma destas metodologias, fundamental para a melhoria da formulação de políticas públicas, é a avaliação de impacto. Segundo a BetterEvaluation¹⁴, uma avaliação de impacto é aquela que forneça informação sobre os efeitos produzidos por uma intervenção – independentemente de

¹⁴ Organização não-governamental cujo objetivo é aprimorar a prática e a teoria de avaliação nas organizações, e que catalogara em sua plataforma informações sobre cerca de 300 métodos e processos de avaliação. Saiba mais em <https://www.betterevaluation.org/>.

serem positivos, negativos, diretos ou indiretos. (PEERSMAN, 2015). Ou seja, trata-se de observar se as causas de uma mudança no cenário avaliado podem ser atribuídas a algum aspecto específico, independentemente dos demais elementos envolvidos, estabelecendo uma relação causal entre ambos. A pergunta básica de avaliação de impacto constitui essencialmente um problema de inferência causal (PEERSMAN, 2015).

Uma avaliação de impacto depende de critérios cientificamente rigorosos, e fatores-chaves podem determinar seu sucesso: recursos disponíveis, desenho da avaliação e qualidade dos dados. A avaliação de impacto de políticas públicas tem se tornado um aspecto cada vez mais relevante, com as organizações multilaterais liderando investimentos, engajamentos e disseminação de conhecimento sobre este método. Um símbolo da crescente relevância do tema foi o Prêmio Nobel de Economia de 2019, concedido a Esther Duflo, Michael Kremer, e Abhijit Banerjee "por sua abordagem experimental para aliviar a pobreza global" (NOBEL OUTREACH AB, 2019).

Contudo, a escassez de recursos dedicados para a avaliação de políticas públicas, como orçamento ou equipe, limita as opções de projetos e de técnicas possíveis de serem implementadas para o avaliador. Consequentemente, avaliações precisam ser combinadas com outras fontes de informações disponíveis, que ajudem a sustentar seu poder de explicação (BAMBERGER; RAO; WOOLCOCK, 2010).

O desenho de uma avaliação de impacto passa pela necessidade de se estabelecer um contrafactual (GERTLER et al., 2018). A forma como o impacto é avaliado envolve criar dois grupos a serem observados separadamente: o grupo de tratamento e o de controle. Quanto melhor a formação dos dois grupos, melhor será a capacidade da avaliação de mensurar os impactos (PEERSMAN, 2015).

O grupo de tratamento é composto pela população, ou uma de suas amostras, que foi alvo da intervenção. Cabe ressaltar que o nome “tratamento” surge da ideia de que a intervenção em questão tenha sido ativamente aplicada na população com o intuito de beneficiá-la, normalmente através uma política pública, mas uma intervenção pode ser não-intencional, quando o termo “grupo de tratamento” pode ser substituído por outro mais intuitivo para o evento analisado; por exemplo “grupo de atingidos”, que será utilizado quando se abordar especificamente o caso do rompimento da barragem de Fundão neste estudo de caso.

Já o grupo de controle é formado por uma população que não recebeu a intervenção, mas que seja o mais similar possível com o grupo de tratamento. Também se trata de um termo que

indica um grupo de pessoas que, mesmo sem ter sido o “tratamento”, foi monitorado pelo avaliador durante todo o processo ou recebido um placebo, o que nem sempre será verdadeiro.

Mas além da boa delimitação dos grupos, os dados que serão avaliados precisam ser da melhor qualidade disponível. Um projeto pode incluir em seu desenho a mensuração de indicadores coletados diretamente com a população de cada grupo, mas uma alternativa é recorrer a dados secundários, como registros administrativos e pesquisas domiciliares generalizadas (GERTLER *et al*, 2018).

Por fim, o bom desenho de uma avaliação de impacto passa também pela escolha de bons indicadores. Ainda que indicadores qualitativos sejam relevantes para ajudar a se criar o contexto e enriquecer conclusões (BAMBERGER *et al*, 2010), sem indicadores objetivos torna-se inviável mensurar e comparar efeitos. Além disso, como o impacto de um programa pode ir além do que pretendia com o seu objetivo, é interessante que os indicadores utilizados abordem mais dimensões do que aquelas em que se haja a expectativa de impacto.

Diante deste contexto, este trabalho visa contribuir para enriquecer políticas públicas, não especificamente olhando para uma delas, mas através da avaliação de uma intervenção não-intencional (o rompimento da barragem de Fundão) na vida de pessoas que, por sua vez, são atualmente alvos de políticas de reparação de danos causados especificamente pelo desastre (FUNDAÇÃO RENOVA, 2019), para as quais estudos desta natureza poderiam ser úteis.

A metodologia deste trabalho será o desenvolvimento de um estudo de caso a partir de uma avaliação de impacto como forma de demonstrar o poder do uso deste tipo de ferramenta para a formação de evidências e o enriquecimento de ideias sobre o combate à pobreza e à vulnerabilidade social, inspirado pela proposta do instituto Abdul Latif Jameel Poverty Action Lab (J-PAL) em seu portal, "*Policy Insights*" (J-PAL, 2019), que cria meta-análises de avaliações de impacto ponderando efeitos de intervenções já existentes para que projetos futuros possam tomar decisões melhor embasadas, talvez encontrando novas soluções ou opções de ações mais efetivas.

Gertler *et al* (2018), com o suporte tanto do Banco Mundial quanto do Banco Interamericano de Desenvolvimento (BID), elaboraram o manual "Avaliação de Impacto na Prática" sobre a aplicação de avaliações de impacto para políticas públicas, disponível gratuitamente em várias línguas, com o objetivo de apoiar planejadores e executores de políticas em todas as áreas, para que consigam aplicar cada vez mais este tipo de análise, enriquecendo o mundo inteiro com maior nível de conhecimento sobre quais ações públicas em quais contextos

são positivas (ou negativas). É um conteúdo extensivo que descreve passo a passo como planejar e executar avaliações de impacto. Nesse sentido, apesar de outras fontes interessantes de avaliação de impacto tenham sido consultadas, como Peersman (2015), Bamberger et al (2010) e Menezes Filho e Pinto (2017), o conteúdo abaixo é majoritariamente baseado no manual de Gertler et al (2018).

O resultado de uma avaliação de impacto pode usar a seguinte expressão. Considerando:

- P a intervenção aplicada em uma população;
- Y o indicador de resultado de interesse;

Temos que:

$$\Delta = (Y|P = 1) - (Y|P = 0)$$

Existem diversos métodos para executar uma avaliação de impacto, que variam de acordo com a qualidade dos dados e as possibilidades de se sustentar pressupostos relevantes. Os principais métodos citados pela literatura são:

Seleção aleatória (RCT). Consiste em gerar duas amostras aleatórias dentro de uma mesma população: uma será o grupo tratamento e a outro o grupo de controle. Estes grupos inicialmente fazem parte de um mesmo público-alvo ou população, portanto, a aleatoriedade permite pressupor com confiança que os grupos será estatisticamente iguais, portanto, qualquer diferença entre eles após a intervenção pode ser atribuída a ela. Para que seja possível fazer uma avaliação por seleção aleatória de qualidade, as técnicas estatísticas aplicadas deverão estar presentes desde a seleção dos participantes, ou seja, os avaliadores precisam ser parte da equipe, ou ao menos prestar-lhes consultorias para conduzir a intervenção de maneira adequada. No caso do rompimento da barragem de Fundão, esse método só seria possível para avaliar o impacto de eventuais programas de reparação da Fundação Renova, caso ela adotasse critérios de seleção aleatória para definir quem participará de tais ações.

Variáveis instrumentais (VI). Na maior parte das vezes, mesmo participando do planejamento de políticas, o avaliador perceberá que é impraticável a seleção aleatória ideal de cada grupo. Por exemplo, quando o ingresso na política é voluntário, universal para toda a população, e passível de desistências dentro dos grupos (alguém selecionado para o tratamento desiste de participar, ou alguém do controle resolve participar). Nesses casos, o avaliador ainda pode se aproveitar de uma variável instrumental exógena à política e que influencie a decisão da

pessoa por participar da ação, desde que esta variável não seja relacionada às características avaliadas, nem esteja sob o controle do avaliado. VI permite avaliar o impacto médio do programa para um subconjunto da população elegível, mas que não necessariamente seria o impacto na população total. Dentro da nossa base de dados, isso seria possível se, por exemplo, tentássemos fazer uma avaliação com famílias participando de políticas da Fundação Renova disponíveis para todos os atingidos que desejassem aderir.

Regressão descontínua (RDD). Se a variável critério para receber o tratamento é contínua e se é possível estabelecer uma linha de corte em algum valor específico da variável, este método permite que o avaliador use a parcela da população com valores próximos à linha para identificar potenciais impactos. RDD pressupõe que a diferença entre os indivíduos, cujo resultado no indicador usado como critério estejam bem próximos da linha, seja superior ou inferior ao valor dela. É fundamental que a linha de corte adotada não sirva para outras formas de intervenção, pois isso comprometeria o poder de explicação da avaliação. O RDD, para estimar o impacto do programa, compara as diferenças entre o resultado médio do grupo quase elegível e do grupo elegível por resultados bem próximos à linha. Imaginando sua aplicação para nosso estudo de caso, poderíamos tentar uma RDD com grupos de pessoas que estão no limite entre a definição de pobreza e de extrema pobreza do CadÚnico para avaliar efeitos mais específicos.

Diferenças em diferenças (DD). Este método é possível quando se possui informações sobre os indicadores avaliados antes e depois da intervenção, para ambos os grupos. A partir dessas informações, o avaliador pode comparar as variações dos dois grupos ao longo do tempo (daí o nome: analisar a diferença das diferenças originais), e analisar se a variação dos indicadores do grupo de tratados foi diferente do que seria esperado na falta da intervenção, baseando-se justamente no que ocorreu com o grupo de controle. Este método é o adotado neste trabalho e será detalhado abaixo.

Pareamento (matching, em inglês). O objetivo do pareamento é elaborar estatisticamente um grupo de controle que seja o mais parecido possível com o grupo tratado. É uma comparação essencialmente estatística, a partir de variáveis que buscam emular um grupo de controle de modo que cada indivíduo possua um "par" no grupo de tratamento.

Por exemplo, no caso do IVF, após identificar as características da nossa população que sofreu a intervenção, deveríamos buscar, em uma outra população, indivíduos com as mesmas características dos atingidos. O pareamento é uma operação complexa e que exige muita precisão,

já que é muito difícil evitar vieses, mas é um modelo importante para os casos em que não há outras opções, o que na prática é a realidade de boa parte das políticas públicas (BAMBERGER *et al*, 2010).

4.2 Avaliação de impacto com o método Diferença em Diferenças

Como descrito acima, existem diversos métodos de avaliação, que possuem níveis de qualidade diferentes. Para Menezes Filho e Pinto (2017) o "padrão-ouro" é o RCT, em que é possível criar as avaliações de impacto com os maiores poderes de explicação, sem a necessidade de pressupostos que vão além das propriedades mais simples da estatística.

Já o modelo DD se torna uma opção interessante quando não é possível realizar os métodos RCT, VI ou RDD, normalmente pela falta de dados de qualidade específicos para avaliar a intervenção. No entanto, a proposta do método é simples e de fácil execução, desde que haja dados históricos ou longitudinais disponíveis. Segundo Gertler *et al* (2018):

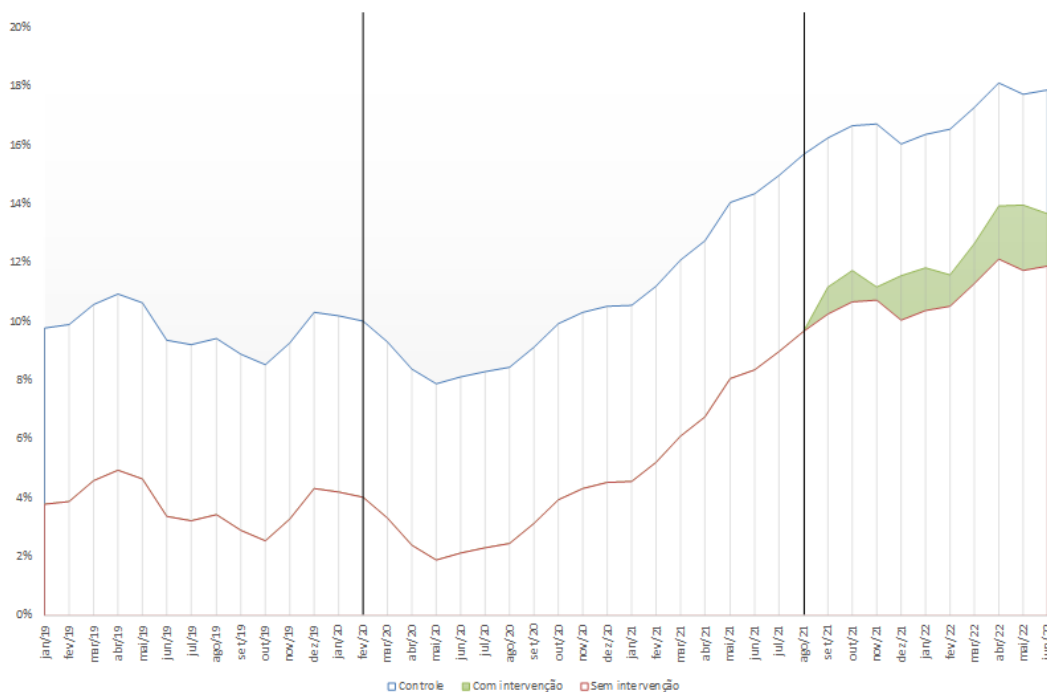
"A diferença em diferenças compara as mudanças nos resultados ao longo do tempo entre os inscritos em um programa (o grupo de tratamento) e os não inscritos (o grupo de comparação). Isso nos permite corrigir quaisquer diferenças entre os grupos de tratamento e de comparação que sejam constantes ao longo do tempo."

Essencialmente, no caso do DD, não existe a necessidade de que o grupo de tratamento e o de controle possuam características iguais. O que precisa ser similar entre os dois grupos é a forma como eles variam no tempo. O pressuposto é: se não houvesse a intervenção, o grupo de tratamento teria uma variação nos indicadores igual àquela observada para o grupo de controle.

Ao assumir este pressuposto, comparando os indicadores de interesse nos períodos anterior e posterior à intervenção, se houve uma alteração na **tendência** do desempenho do indicador no grupo de tratamento, pode-se atribuir este "choque" à intervenção analisada. A

Figura 4-1 pode ser útil para visualizar um exemplo ideal.

Figura 4-1 – Simulação de uma comparação entre as trajetórias do grupo de controle, do grupo de tratamento, com e sem a intervenção



Neste gráfico, temos a simulação de uma intervenção qualquer que usa um indicador em percentual (por exemplo: o percentual de jovens que gostam de ler). O grupo de controle é a linha azul, enquanto o grupo de tratamento é a linha vermelha.

O primeiro ponto a se notar é que os dois grupos não são iguais: o grupo azul sempre teve mais jovens interessados em leitura que o vermelho. Porém, uma segunda observação importante é que a variação, ou a tendência ao longo do tempo, de ambos os grupos é basicamente idêntica: quando o grupo azul aumenta o interesse por leitura, o vermelho também aumenta, e vice-versa.

Suponha que seja executado um programa de incentivo à leitura para o grupo vermelho, a partir de março de 2020, e que dure até agosto de 2021 (as duas linhas pretas verticais na

Figura 4-1). Se o programa não funcionasse, e não tivesse nenhum efeito especial, o grupo vermelho seguiria seguindo o trajeto com as mesmas tendências do azul, conforme a continuação da linha vermelha da figura mostra.

Por outro lado, com o programa sendo bem-sucedido, aumentando o número de leitores graças a sua intervenção, a área verde do gráfico mostra o impacto benéfico que ele causou no grupo vermelho. O importante a se observar é que, graças ao programa, o grupo vermelho traça uma nova trajetória, diferente daquela que seria prevista, fazendo com que ele evoluía em um ritmo mais acelerado do que o grupo azul. Isto implica que o método DD consegue avaliar o impacto a

partir do momento em que ele identifica alterações nas tendências do grupo tratado, quando comparado com o grupo de controle, sejam positivas ou negativas.

Contudo, ele exige um pressuposto forte: apesar do valor dos indicadores não precisar ser o mesmo entre os dois grupos no início da intervenção, suas tendências de evolução precisam ser as mais próximas possíveis na ausência do tratamento. Se outros fatores causarem mudanças de tendência diferentes entre os grupos, isto não será detectável na avaliação por diferenças em diferenças.

Uma forma de mitigar estes riscos é utilizado métodos mistos de avaliação, que podem ajudar a identificar qualitativamente a existência de outros eventos inesperados acontecendo nos grupos (BAMBERGER *et al*, 2010). Mas este tipo de iniciativa infelizmente precisou ser descartado deste trabalho. Assume-se que a seleção dos municípios de controle, conforme gerada pelo método k-NN descrito acima, tenha sido boa o suficiente para que esta não seja uma preocupação, e que as trajetórias do IVF em ambos os grupos possuiriam as mesmas tendências sem o rompimento da barragem de Fundão.

4.2.1 Passo a passo para realizar uma avaliação de impacto DD:

Em geral, os cálculos da avaliação por DD, por si só, são simples: uma sequência de cálculos de subtração. Os maiores desafios residem na estruturação dos dados para realizar a análise. A seguir temos as principais etapas deste processo, conforme estabelecidas pelo Manual de Avaliação de Impacto na Prática de Gertler *et al* (2018).

Formação do grupo de controle. Usar o método de diferenças em diferenças é necessário quando não foi possível acompanhar a intervenção *ex-ante*, por isso será necessário criar seus próprios grupos de tratamento e de controle sinteticamente. Como dito anteriormente, eles não precisam ter características iguais, mas precisam ter tendências de variação o mais próximas possíveis.

Coleta de dados. Para ambos os grupos, o avaliador precisa coletar os valores dos indicadores de interesse em datas anteriores e após o processo de intervenção.

Primeira diferença: Significa calcular, para o grupo de tratamento, a variação no indicador antes e depois da intervenção. Esta informação por si só é chamada por Gertler *et al* (2018) de um

caso de estimativa falsa de contrafactual, não servindo para medir o impacto sem o cálculo da segunda diferença. Considerando:

- A o grupo de atingidos pela intervenção (grupo de tratamento);
- A_1 o valor do indicador para o grupo de atingidos em uma data anterior à intervenção;
- A_2 o valor do mesmo indicador para o mesmo grupo, em uma data posterior à intervenção;

Teremos que a primeira diferença equivale a:

$$\Delta A = A_2 - A_1$$

Segunda diferença: Implica em fazer o mesmo cálculo da diferença entre antes e depois, mas para o grupo de controle. Seguindo o mesmo raciocínio da primeira diferença, temos:

- C_1 o valor do indicador para o grupo de controle em uma data anterior à intervenção;
- C_2 o valor do mesmo indicador para o mesmo grupo, em uma data posterior à intervenção que fora aplicada no grupo de tratamento;

Assim a segunda diferença será igual a:

$$\Delta C = C_2 - C_1$$

Diferenças em diferenças: Por fim, com as variações do tratamento e do controle em mãos, o impacto será medido subtraindo a primeira diferença da segunda:

$$\Delta Y = \Delta A - \Delta C = (A_2 - A_1) - (C_2 - C_1)$$

Uma outra sugestão de Gertler *et al* (2018) sobre como calcular o método DD é através das diferenças entre grupos antes da intervenção, e em seguida a mesma subtração após a intervenção. Assim, teremos a seguinte expressão

$$\Delta Y = \Delta A - \Delta C = (A_1 - C_1) - (A_2 - C_2)$$

O método diferenças em diferenças é interessante porque controla vieses e variáveis que não são observados ao longo do tempo (assumindo seus pressupostos). Como ele traz uma visão das tendências, costuma ser bastante útil para o planejamento de políticas públicas (GERTLER *et al*, 2018).

4.3 A aplicação da avaliação de impacto neste estudo de caso

A avaliação de impacto padrão que foi realizada neste trabalho utilizou o IVF como seu indicador (Y), os anos de 2012 a 2014 como os períodos antes do desastre (a intervenção), e os anos de 2015 a 2018 como os períodos em que os impactos afetam os atingidos.

Dada a polaridade menor-melhor do IVF e de todas as suas variáveis e, considerando que as diferenças são calculadas com o grupo de atingidos sendo o minuendo e com o grupo de controle sendo o subtraendo ($\Delta A - \Delta C$), quando a variação analisada é **positiva**, significa que os municípios atingidos se tornaram **mais vulneráveis** do que os municípios do grupo de controle.

Nesse sentido, é importante perceber a diferença entre usar expressões como “aumentar” ou “diminuir” de expressões como “melhorar” ou “piorar”, pois elas podem ter sentidos opostos no contexto. Para facilitar a compreensão das análises, quando necessário, priorizou-se manter termos repetitivos, porém mais claros, do que preservar a estética de um texto que não seja repetitivo.

4.3.1 Comparações anuais

Estas são as comparações dos valores do indicador no dia 31 de dezembro de cada ano, conforme a base desidentificada do Cadastro Único. Isto implica na possibilidade de contrastar a diferença entre o valor de cada variável dentre os dois grupos:

$$\Delta Y_t = A_t - C_t$$

Esta análise tem como proposta realizar apenas uma descritiva simples entre os valores anuais do IVF e suas componentes, ressaltando que limitar a avaliação a estes valores implicaria em gerar falsos contrafactuais.

4.3.2 Comparações por períodos

Os períodos estão aqui definidos como uma forma de avaliar o impacto do desastre em diferentes horizontes temporais. Gertler *et al* (2018) sugere que o avaliador sequencie uma cadeia resultados, que incluem a entrada de insumos; a execução das atividades; os produtos (os bens materiais e os serviços entregues pela intervenção); os “resultados” alcançados no curto ou médio prazos; e os “resultados finais” em comparação com os objetivos do projeto.

Contudo, estes critérios pressupõem que as intervenções sejam políticas públicas ativas e diretas, aplicadas a partir de uma teoria de mudança. No nosso caso, estamos analisando um evento incidental e súbito e seus efeitos indiretos na população atingida.

Para tanto, será definido um critério de análise similar ao adotado pela organização Agência dos Estados Unidos para o Desenvolvimento Internacional (USAID), através de sua política de avaliação de impactos de longo prazo (ERIE, 2018):

“The difference between a traditional impact evaluation and long-term impact evaluation is that LTIEs measure outcomes over a longer time horizon. USAID program cycles are often three to five years. When impact evaluations are incorporated into the program, they typically measure outcomes that materialize during program implementation, or immediately after. In contrast, LTIEs would measure the impact on the same or other outcomes well after implementation of the program has ended. This type of evaluation provides critical information on whether observed changes are sustained over time, or whether other anticipated, long-term, outcomes result from the program.” (Erie, 2018)

Esta avaliação a avaliação assumirá os resultados de dezembro de 2014 como os últimos disponíveis antes do rompimento da barragem de Fundão, uma vez que a próxima informação se refere à 31 de dezembro de 2015, ou seja, cerca de um mês após o evento.

Optou-se por chamar de "curto prazo" a variação entre os anos de 2014 e 2015. Já quando a avaliação focar em períodos de médio prazo, serão utilizados horizontes de dois ou três anos após o desastre, ou seja, a variação entre os anos de 2014 e 2016 ou entre 2014 e 2017. Por fim, impactos após quatro anos (a variação entre 2014 e 2018) serão considerados de longo prazo, ou seja, a variação de 2014 a 2018. Para registrar em expressão, considerando o ano de 2014 como t_0 , e qualquer outro ano como t_n , temos que:

$$\Delta Y_{tn} = \Delta A_{tn-t_0} - \Delta C_{tn-t_0}$$

4.3.3 Relacionamento entre dimensões demográficas e de ocupação

A principal preocupação deste estudo está em observar os efeitos da composição familiar nas condições de ocupação e de pobreza dos atingidos pelo rompimento da barragem do fundão.

O IVF, conforme descrito acima, possui seis dimensões, e dentre estas 40 variáveis binárias. Apesar de o indicador permitir análises em mais dimensões, sobre assuntos como educação e habitação, o foco deste estudo está em verificar a relação entre a dimensão de vulnerabilidade demográfica com o acesso ao trabalho e a escassez de recursos.

Das 10 variáveis que compõem a vulnerabilidade demográfica do IVF, foram escolhidas quatro de interesse:

- V₁. Alguma mulher teve filho nascido vivo no último ano
- V₆. Presença de idoso;
- V₇. Ausência de cônjuge; e
- V₈. Menos da metade dos membros encontram-se em idade ativa.

Conforme demonstrado por Barros, Carvalho e Franco (2006), esta é uma das vantagens de o IVF ser um indicador escalar desagregável: é possível decompor as populações em grupos demográficos específicos. Neste caso, o interessante é calcular separadamente os impactos do desastre nas famílias de toda a população; apenas nas famílias consideradas vulneráveis para cada uma destas variáveis (valor igual a 1) e nas famílias consideradas resilientes naquela mesma variável.

Ou seja, para cada uma destas variáveis, foram elaborados:

- (a) um primeiro subgrupo de atingidos, incluindo apenas as pessoas consideradas vulneráveis sob a perspectiva daquela característica;
- (b) um segundo subgrupo de vulneráveis sob o mesmo critério, mas a partir da população do grupo de controle;
- (c) um terceiro subgrupo de atingidos, agora incluindo apenas as pessoas consideradas resilientes sob a perspectiva daquela mesma característica;
- (d) um quarto subgrupo, com os mesmos critérios, mas para a população do grupo de controle.

Concluindo, cada variável demográfica gerou seus próprios grupos de resilientes e de vulneráveis, dentro dos grupos de tratamento e do grupo de controle, gerando assim um universo com 18 subgrupos diferentes.

Uma vez definidos quais são os grupos avaliados, priorizou-se focar nos efeitos do desastre no IVF global, na dimensão de acesso ao trabalho e em cada uma de suas seis variáveis, bem como a dimensão de escassez de renda e suas três variáveis (veja a Tabela 4-1).

Tabela 4-1 – Indicadores e variáveis de acesso ao trabalho e escassez de recursos a serem avaliadas

Índice de Vulnerabilidade das Famílias (IVF)	Índice da dimensão de acesso ao trabalho (T)	T1. Menos da metade dos membros em idade ativa encontram-se ocupados
		T3. Ausência de ocupado no setor formal
		T4. Ausência de ocupado em atividade não-agrícola
		T5. Ausência de ocupado com rendimento superior a um salário-mínimo
		T6. Ausência de ocupado com rendimento superior a dois salários-mínimos
		Índice da dimensão de escassez de renda (R)
	R2. Renda familiar per capita inferior à linha de pobreza; e	
	R3. Maior parte da renda familiar advém de transferências.	

4.4 Seleção do grupo de tratamento

Dada a condição *ex-post* desta avaliação mensurada anos após o desastre, houve a necessidade de elencar grupos de tratamento e de controle com técnicas indiretas.

Para definir quais municípios foram atingidos, optou-se pelo critério registrado do Termo de Transação e Ajustamento de Conduta (TTAC), assinado em 2016 pela empresa Samarco Mineração S.A., e seus acionistas Vale S.A. e BHP Billiton Brasil LTDA.; em conjunto com instituições do governo federal e dos estados de Minas Gerais e Espírito Santo (BRASIL, 2016). Este documento definiu quais municípios foram considerados atingidos pelo rompimento da barragem do Fundão e formou as bases da Fundação Renova, que por sua vez considera estes os

municípios a serem atendidos pelos seus programas, por exemplo os investimentos para projetos de saneamento (FUNDAÇÃO RENOVA, 2018)¹⁵:

Mariana, Governador Valadares, Barra Longa, Rio Doce, Santa Cruz do Escalvado, Sem-Peixe, Rio Casca, São Pedro dos Ferros, São Domingos do Prata, São José do Goiabal, Raul Soares, Dionísio, Córrego Novo, Pingo d'Água, Marliéria, Bom Jesus do Galho, Timóteo, Caratinga, Ipatinga, Santana do Paraíso, Ipaba, Belo Oriente, Bugre, Iapu, Naque, Periquito, Sobrália, Fernandes Tourinho, Alpercata, Governador Valadares, Tumiritinga, Galileia, Conselheiro Pena, Resplendor, Itueta e Aimorés; além dos capixabas Baixo Guandu, Colatina, Marilândia e Linhares.

São municípios muito heterogêneos, e em vários casos com baixo dinamismo econômico (MINAS GERAIS, 2016). Esse será um ponto bastante relevante, embora difícil de se destacar na avaliação: para que fosse possível elaborar um grupo de controle mais qualificado, foi necessário manter o grupo de atingidos (tratamento) como uma região única. Por exemplo, não foi possível criar um grupo de controle adequado apenas para os quatro municípios da microrregião afetada diretamente pela lama do rompimento (MINAS GERAIS; 2016).

4.5 Seleção do grupo de controle

Para elencar os municípios de controle contrafactual da avaliação do impacto, optou-se por pela mesma técnica utilizada no trabalho de Rezende (2021) para a sua avaliação de impacto do rompimento nos fluxos migratórios, o *K-Nearest Neighbor* (k-NN), descrita a seguir.

4.5.1 O algoritmo *K-Nearest Neighbor* (k-NN)

O método de k Vizinhos Mais Próximos, conhecido como k-NN (do inglês, *k Nearest Neighbors*), é um algoritmo de aprendizado de máquina para classificações e regressões (DIHANSTER, 2020).

¹⁵ Atualmente, o Ministério Público Federal vem protagonizando mudanças estruturais nos acordos celebrados entre a Fundação Renova, os entes da federação e grupos sociais atingidos, o que pode ser acompanhado em seu portal Caso Samarco: <http://www.mpf.mp.br/grandes-casos/caso-samarco> (BRASIL, 2022b).

Ele visa qualificar um elemento a partir das características de outros elementos que estejam próximos a ele. Esta proximidade é definida através de parâmetros que distribuam os elementos em um espaço n-dimensional. De acordo com a posição de todas as observações neste espaço, é possível classificar um dado ponto X analisando qual é a classe predominante dentre seus “vizinhos” (os K pontos mais parecidos o possível com o ponto X). Trata-se de um algoritmo eminentemente baseado em Distâncias Euclidianas, mas ele aceita várias outras medidas de distância alternativas (DIHANSTER, 2020).

O k-NN se caracteriza por ser um algoritmo supervisionado de aprendizagem preguiçosa. Isto significa que ele se utiliza de um banco de dados de treinamento e, a partir de parâmetros importados deste banco, infere uma função de classificação para ser aplicada em outras bases de dados importadas (IBM CLOUD EDUCATION, [2020]). Além disso, ele não aprende a partir de uma função discriminativa, como regressões logísticas, mas identifica o conjunto de dados de treinamento no momento da própria análise, o que pode aumentar muito o tempo de processamento para o aprendizado, uma vez que cada ponto precisa ser comparado com todos os demais em cada iteração (RASCHKA, [2016]). Segundo a IBM Cloud Education ([2020]):

"Its ease of use and low calculation time make it a preferred algorithm by data scientists, but as the test dataset grows, the processing time lengthens, making it less appealing for classification tasks. KNN is typically used for recommendation engines and image recognition."

Além disso, ele também é considerado não-paramétrico, o que neste caso implica que não são necessários pressupostos para a distribuição dos dados, características que nos é relevante, dado que não houve nenhum critério específico para a seleção dos municípios (DATACAMP, [2018]).

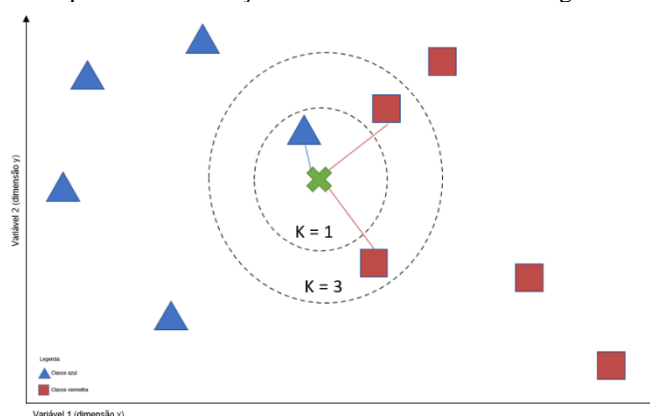
Este foi o método utilizado por Rezende (2021) em sua estratégia de identificação dos municípios de controle para a sua avaliação do impacto do rompimento da barragem de Fundão na migração dos municípios atingidos, o que ajuda a aproximar bastante os resultados, ainda que em variáveis diferentes, para eventuais meta-análises sobre o desastre.

Para definir o algoritmo, três fatores devem ser estipulados (FERRERO, 2009): (1) o número de observações da base de dados destinada ao treinamento; (2) as variáveis de parâmetros que servirão de critério para a classificação, e (3) o valor de K, ou seja, o número de vizinhos a serem computados dentro do espaço.

Uma boa delimitação das variáveis que servirão como parâmetros é fundamental, pois é a partir dos dados encontrados nela que o k -NN conseguirá distribuir as observações em um espaço dimensional adequado. Quanto menor a diferença entre valor do ponto X em uma dada variável e o valor dos outros pontos, menores serão as distâncias entre eles na enésima dimensão que apresenta esta variável.

Além disso, o valor de k é quem vai estabelecer o número de vizinhos a ser considerado. Podemos perceber, utilizando um exemplo clássico com duas dimensões (variáveis), como o valor de k é utilizado pelo algoritmo. Temos na Figura 4-2 uma série de pontos classificados de forma binária (vermelho quadrado ou azul triangular). Distribuídas por todo o espaço, há onze observações: cinco vermelhas, cinco azuis, e o ponto X que queremos classificar.

Figura 4-2 – Exemplo de classificação do método K-Nearest Neighbor com $k = 1$ e $k = 3$



Dada a distribuição no gráfico, com um $k = 1$, o único ponto analisado, e o mais próximo de X, é um ponto *azul triangular*. Portanto, o k -NN classificará o ponto X como *azul triangular*. Por outro lado, com um $k = 3$, percebe-se que a predominância dos pontos mais próximos inverte para a classe *vermelho quadrado*, ou seja, mesmo que um único *ponto azul* seja literalmente a menor distância euclidiana do ponto X, quando se amplia para buscar os três vizinhos mais próximos, a classificação que prevalece para o ponto X torna-se a *vermelho quadrado*.

Percebe-se o quanto que o valor de K é tão relevante quanto as posições de cada ponto no espaço. Assim, conclui Rezende (2021) que, para avaliações como esta, como não há um único valor de k que se adeque, é preciso estabelecer um valor adequado para este problema.

Uma vez aplicado o método k-NN, tem-se em mãos uma ferramenta capaz de predizer ou classificar cada observação de uma base de dados (a base de dados de teste). Para medir a qualidade desta predição, uma ferramenta utilizada é a matriz de confusão¹⁶.

Esta é uma matriz que compara resultados observados com resultados preditos pelo modelo. Dada uma classificação binária e definido um resultado como "positivo" e outro como "negativo", define-se como "falso positivo" o resultado que classifique incorretamente um ponto como positivo, e "falso negativo" a predição incorreta de um ponto como negativo. A matriz de confusão é aquela que enumera os casos de acertos e de erros da predição, como no modelo da Tabela 4-2.

Tabela 4-2 – Modelo de matriz de confusão

Predições	Referências	
	Positivos	Negativos
Positivos	A	B
Negativos	C	D

A partir da matriz é possível mensurar a acurácia diagnóstica (BISSOLI JUNIOR; DIEHL; PRADO, 2020), uma medida da capacidade de um teste de identificar corretamente a classe dos dados observados. Define-se como “sensibilidade” a medida do percentual de observações em que o teste prediz corretamente que um ponto X é positivo, e “especificidade” o percentual de predições corretas quando o ponto X é negativo.

Portanto, a partir da acurácia diagnóstico, será possível estimar a qualidade do aprendizado do algoritmo e de sua aplicação, o que será fundamental para a nossa estratégia de avaliação.

4.5.2 Variáveis de pareamento para o algoritmo k-NN

O grupo de controle foi formado a partir de características estruturais dos municípios para lidar com a vulnerabilidade social. Ele deveria possuir não apenas populações com composições familiares parecidas, mas também estruturas de serviços públicos similares. Entende-se que indicadores agregados para a população total do município são melhores do que considerar apenas

¹⁶ Exemplos de uso da matriz de confusão como modelo de análise de acurácia diagnóstica do kNN podem ser conferidos em Alam (2022), Beena (2020) e Gurucharan (2020).

os cadastrados no CadÚnico, porque isso ajudaria a considerar as parcelas da população que tenham chances de entrada e de saída das famílias cadastradas, alterando sua condição de pobreza ao longo dos anos avaliados.

Para tanto, dividiu-se a lista de indicadores em duas grandes fontes: o Censo 2010, para calcular as médias municipais das variáveis que compõem as dimensões de vulnerabilidade demográfica, acesso ao trabalho e escassez de renda; e o IMRS 2014, com indicadores representativos da estrutura de prestação de serviços assistenciais: saúde, educação e assistência social. O IMRS também forneceu dois dados sobre o município como um todo: o total de habitantes estimados para 2014, e sua taxa de urbanização. Abaixo seguem as descrições fornecidas Fundação João Pinheiro ([2021]), segundo seu dicionário de variáveis do IMRS.

População total (estimativas ajustadas). População total residente, utilizada para calcular todos os indicadores da base de dados do IMRS. Corresponde às estimativas populacionais produzidas pela Fundação João Pinheiro desde 2011 para o IMRS. Em 2018, o IBGE realizou uma revisão de suas estimativas populacionais anteriores e os resultados apresentados são muito próximos aos das estimativas da FJP. Optou-se por manter os resultados das estimativas da FJP para permitir que os usuários possam obter resultados semelhantes aos nossos neste banco de dados. As novas estimativas do IBGE podem ser obtidas em: <https://www.ibge.gov.br/estatisticas/sociais/populacao/9103-estimativas-de-populacao.html>.

Fontes: IBGE/Fundação João Pinheiro.

Taxa de urbanização. Razão entre o número total de pessoas residentes na área urbana do município e a sua população residente total. Para os anos de 2000 e 2010, os dados de população são censitários. Para os anos intercensitários, a população foi estimada por interpolação. *Fontes:* IBGE/Fundação João Pinheiro.

Taxa de atendimento da educação básica. Indicador de acesso das crianças e jovens à educação básica. É a razão entre o número de matrículas de crianças e jovens de 4 a 17 anos de idade e o número total de pessoas nesta faixa etária, multiplicada por 100. *Fontes:* a) Matrículas: Ministério da Educação/INEP; b) População: IBGE/Fundação João Pinheiro.

Proporção da população atendida pela Estratégia de Saúde da Família¹⁷. Razão entre a capacidade de atendimento e a população total do município. A capacidade de atendimento corresponde ao produto do número médio de equipes no ano (soma dos números de equipes de saúde da família em cada mês do ano, dividido por 12) e da estimativa de atendimento médio por equipe, de acordo com a SES-MG. *Fontes*: a) Equipes do PSF: Secretaria de Estado de Saúde de Minas Gerais (SES-MG); b) população: Estimada pela Fundação João Pinheiro.

Índice de Gestão Descentralizada do Sistema Único de Assistência Social. Indicador que avalia a gestão descentralizada dos serviços, programas, projetos e benefícios socioassistenciais na esfera dos municípios. Corresponde ao resultado da média aritmética ponderada do Indicador de Desenvolvimento de Centros de Referência da Assistência Social (IDCRAS), com peso 4, e da execução financeira ajustada, com peso 1 (vide definições do IDCRAS e do Indicador Execução Financeira Ajustada¹⁸). Varia de 0 (zero) a 1 (um). Quanto mais próximo de 1, melhor é o desempenho da gestão e maior será o valor de apoio financeiro repassado aos municípios como forma de aprimoramento da gestão¹⁹. *Fontes*: Ministério da Cidadania e Matriz de informação Social – MI Social.

¹⁷ A Estratégia de Saúde da Família é um plano "de expansão, qualificação e consolidação da atenção básica por favorecer uma reorientação do processo de trabalho com maior potencial de aprofundar os princípios, diretrizes e fundamentos da atenção básica, de ampliar a resolutividade e impacto na situação de saúde das pessoas e coletividades, além de propiciar uma importante relação custo-efetividade" (Brasil, 2012a).

¹⁸ Disponíveis no mesmo dicionário de dados do IMRS, disponibilizado pela Fundação João Pinheiro ([2021])

¹⁹ Com a inclusão deste indicador, espera-se que haja um reforço do pressuposto de que os eventuais erros de registros administrativos comentados na descrição da base de dados do Cadastro Único sejam aleatoriamente bem distribuídos dentre os grupos de municípios atingidos e de controle.

Tabela 4-3 – Relação de indicadores para a seleção de municípios para o grupo de controle através do método de aprendizagem *k-NN*

Censo 2010	IMRS 2014
Presença de dependente (0 a 1 ano)	População total (estimativas ajustadas)
Presença de dependente (0 a 1 anos)	Taxa de urbanização
Presença de dependente (0 a 12 anos)	Taxa de atendimento da educação básica
Presença de dependente (0 a 17 anos)	Proporção da população atendida pela Estratégia de Saúde da Família
Presença de dependente (0 a 21 anos)	Índice de Gestão Descentralizada do Sistema Único de Assistência Social
Presença de dependente (65 ou mais anos)	
Ausência de cônjuge	
Presença de deficiente	
Presença de filho com mãe ausente	
Minoria está em idade ativa	
Minoria é ocupada	
Ausência de ocupado formal	
Ausência de ocupado não agricultor	
Ausência de ocupado com mais de 1 salário	
Ausência de ocupado com mais de 2 salários	
Domicílio extremamente pobre	
Domicílio pobre	
Minoria da fonte de renda é por trabalho	

Fonte: Censo 2010 e IMRS 2014.

4.5.3 Estratégia de implementação do algoritmo *k-NN*

O objetivo da implementação do *k-NN* é elencar, no estado de Minas Gerais²⁰, os municípios que formarão um grupo de controle como contrafactual ao grupo de atingidos pelo desastre, considerando como critério de pareamento as características descritas acima e resumidas na Tabela 4-3.

São imputadas na análise as 23 variáveis de pareamento, devidamente escalonadas²¹ para evitar vieses de peso por dimensionamentos, mais uma 24^a variável binária, chamada “Atingido”: definindo com valor 1 cada município atingido pelo desastre e 0 para os não atingidos.

²⁰ Diferente de Rezende (2021), que incluiu os municípios do Espírito Santo que também foram atingidos para a formação de seus grupos de tratamento e controle na avaliação de impacto do desastre nas migrações.

²¹ Segundo Navlani (2018): “*KNN requires scaling of data because KNN uses the Euclidean distance between two data points to find nearest neighbors. Euclidean distance is sensitive to magnitudes. The features with high magnitudes will weight more than features with low magnitudes.*”

Todos os municípios foram separados em duas amostras por sorteio probabilístico via amostragem aleatória simples sem reposição (AASS): a amostra de teste e a amostra de treinamento. Ao final, obtiveram-se 427 observações para teste (incluindo 17 municípios do grupo de atingidos), enquanto a amostra de teste recebeu as demais 426 (18 atingidos).

Através dos valores nos parâmetros da amostra de treinamento, o algoritmo tentou prever, para a amostra de teste, se cada município foi classificado como atingido ou como não-atingido. Ou seja, o algoritmo compara cada observação da amostra de teste com os parâmetros dos municípios da amostra de treinamento e, utilizando-se destas informações, infere a condição de atingido nos municípios de teste. Quando o algoritmo eventualmente gera um falso positivo, significa que ele encontrou um município não atingido no restante do estado com características tão semelhantes aos 35 atingidos que se concluiu, por engano, que este seria um dos 35. O fato de que os falsos positivos ocorrem justamente por serem parecidos com os casos reais nos parâmetros que nos interessam é um indício de que eles podem ser elencados para o grupo de controle.

Nesse sentido, o k - NN precisa ser criado a partir de uma modelagem bem equilibrada, com uma acurácia que possua uma sensibilidade baixa a ponto de gerar um número razoável de falsos positivos, mas mantendo uma especificidade alta, o que demonstraria como os parâmetros estão sendo capazes de encontrar os municípios efetivamente atingidos.

4.5.4 Municípios elencados para o Grupo de Controle

A aprendizagem pelo algoritmo k - NN foi programada para ser executada em três rodadas principais.

Na primeira delas, foram geradas dez previsões classificatórias pelo k - NN , em que cada iteração possuía um valor de K entre 1 e 10. Para cada K , o aprendizado gerado a partir da amostra de treinamento buscou prever quais seriam os municípios da amostra de teste que foram atingidos. Com estes dados foram gerados os indicadores de acurácia de cada K : percentual de erro, sensibilidade e especificidade.

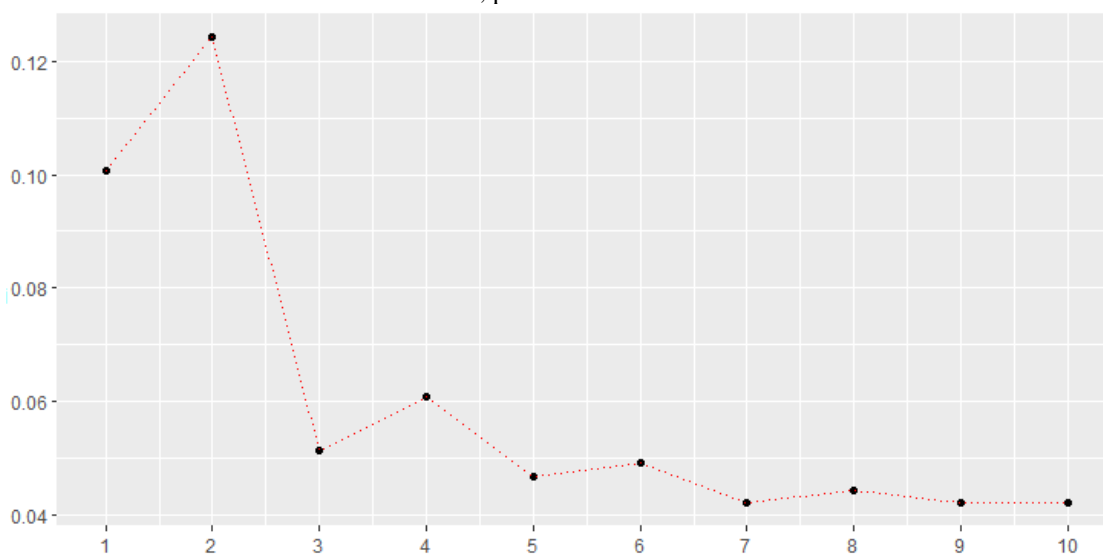
Tabela 4-4 – Percentual de erro, sensibilidade e especificidade pela técnica k-NN, por valor de K

K	Nível de erro	Sensibilidade	Especificidade
1	9,80%	93%	18%
2	13,30%	90%	35%
3	7,30%	98%	18%
4	8,20%	97%	18%
5	5,40%	99%	0%
6	4,90%	99%	0%
7	4,90%	100%	0%
8	4,40%	100%	0%
9	4,00%	100%	0%
10	4,20%	100%	0%

Fontes: Censo 2010 e IMRS 2014.

Na Figura 4-3 é possível observar que os maiores percentuais de erro ocorrem quando o valor de K é mais baixo, com destaque para o $K = 2$. A partir de $K = 5$, os percentuais de erro são tão baixos que tornam nossas chances de encontrar falsos positivos inviáveis.

Figura 4-3 – Percentual de erro na capacidade de previsão de municípios atingidos na amostra de teste pela técnica de k-NN, por valor de K

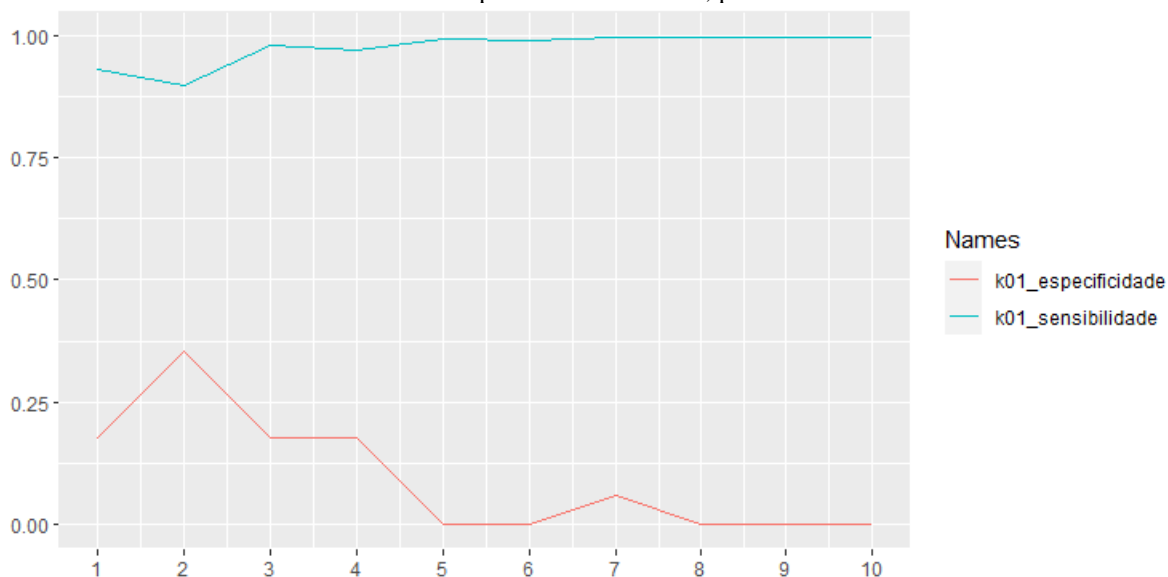


Fontes: Censo 2010 e IMRS 2014.

Analisando especificamente os níveis de sensibilidade e de especificidade gerados (Figura 4-4), é possível identificar que a iteração com valor de $K = 2$ gera uma sensibilidade de 90%,

mantendo uma especificidade de 35%. Considerando os objetivos da análise, optou-se por definir o valor de $K = 2$ para as próximas rodadas.

Figura 4-4 – Níveis de sensibilidade e de especificidade na capacidade de previsão de municípios atingidos na amostra de teste pela técnica de k -NN, por valor de K



Fontes: Censo 2010 e IMRS 2014.

Na segunda rodada de k -NN, a partir do valor de $K = 2$, foram realizadas 1.000 iterações de aprendizado pela amostra de treinamento para a identificação de atingidos na amostra de teste, identificando seus respectivos níveis de acurácia.

O objetivo fora identificar, dentre as mil iterações, as cinco com os maiores níveis de especificidade e, dentre estas, selecionar aquela que possuía a menor sensibilidade. A Tabela 4-5 destaca estes casos, identificando seus respectivos níveis de sensibilidade e de especificidade.

Tabela 4-5 – Níveis de sensibilidade e especificidade para as cinco iterações de maior especificidade na segunda rodada de k -NN

Iteração	Sensibilidade	Especificidade
398	91%	59%
299	90%	53%
412	88%	47%
657	87%	41%
615	86%	35%

Fontes: Censo 2010 e IMRS 2014.

Dentre elas, seguindo a estratégia de criação do grupo de controle, optou-se por escolher a 615ª iteração, pois esta apresenta uma sensibilidade de 86%, ou seja, a que apresenta maiores chances de gerar falsos positivos ainda bem similares aos reais municípios atingidos. Esta iteração gerou as predições descritas na Tabela 4-6.

A estes 56 municípios será atribuída a classificação na variável “atingido” de um valor igual a 1, ou seja, a partir de agora a análise deverá simular que os falsos positivos da amostra de teste sejam também parte do grupo de atingidos. Com isto, teremos uma lista com parâmetros reforçados para a aprendizagem do *k-NN* na próxima etapa.

Tabela 4-6 – Comparação entre o real número de atingidos e não atingidos (referência), com suas respectivas predições *k-NN*, para a base de dados de teste

Predições	Referências	
	Não atingidos	Atingidos
Não atingidos	354	11 ¹
Atingidos	56 ²	6

Notas: (1) Falsos positivos: os municípios com potencial de participar do grupo de controle. (2) Falsos negativos. Fontes: Censo 2010 e IMRS 2014.

Na terceira rodada, há uma inversão das bases de teste e de treinamento, a fim de encontrar na outra amostra de 410 municípios não atingidos, quais destes poderiam ser considerados como grupo de controle. Assim, uma nova sequência de mil iterações de aprendizagem de *k-NN* foi executada, mas desta vez as amostras foram invertidas: a aprendizagem será com a amostra de testes, enquanto a classificação será para a amostra de treinamento. Os resultados de acurácia diagnóstica encontrados, para as iterações com os cinco maiores níveis de especificidades, estão na Tabela 4-7.

Tabela 4-7 – Níveis de sensibilidade e especificidade para as cinco iterações de maior especificidade na terceira rodada de *k-NN*

Iteração	Sensibilidade	Especificidade
84	73,5%	88,9%
49	74,5%	83,3%
544	72,8%	77,8%
500	72,3%	72,2%
812	73,0%	66,7%

Fontes: Censo 2010 e IMRS 2014.

Desta forma optou-se pela 500^a iteração, cuja relação entre as referências (municípios realmente atingidos ou não) e as predições da iteração podem ser observadas na Tabela 4-8.

Tabela 4-8 – Comparação entre o real número de atingidos e não atingidos (referência), com suas respectivas predições k-NN, para a base de dados de treinamento

Predições	Referências	
	Não atingidos	Atingidos
Não atingidos	295	5 ²
Atingidos	113 ¹	13

Notas: (1) Falsos positivos: os municípios com potencial de participar do grupo de controle. (2) Falsos negativos.
Fontes: Censo 2010 e IMRS 2014.

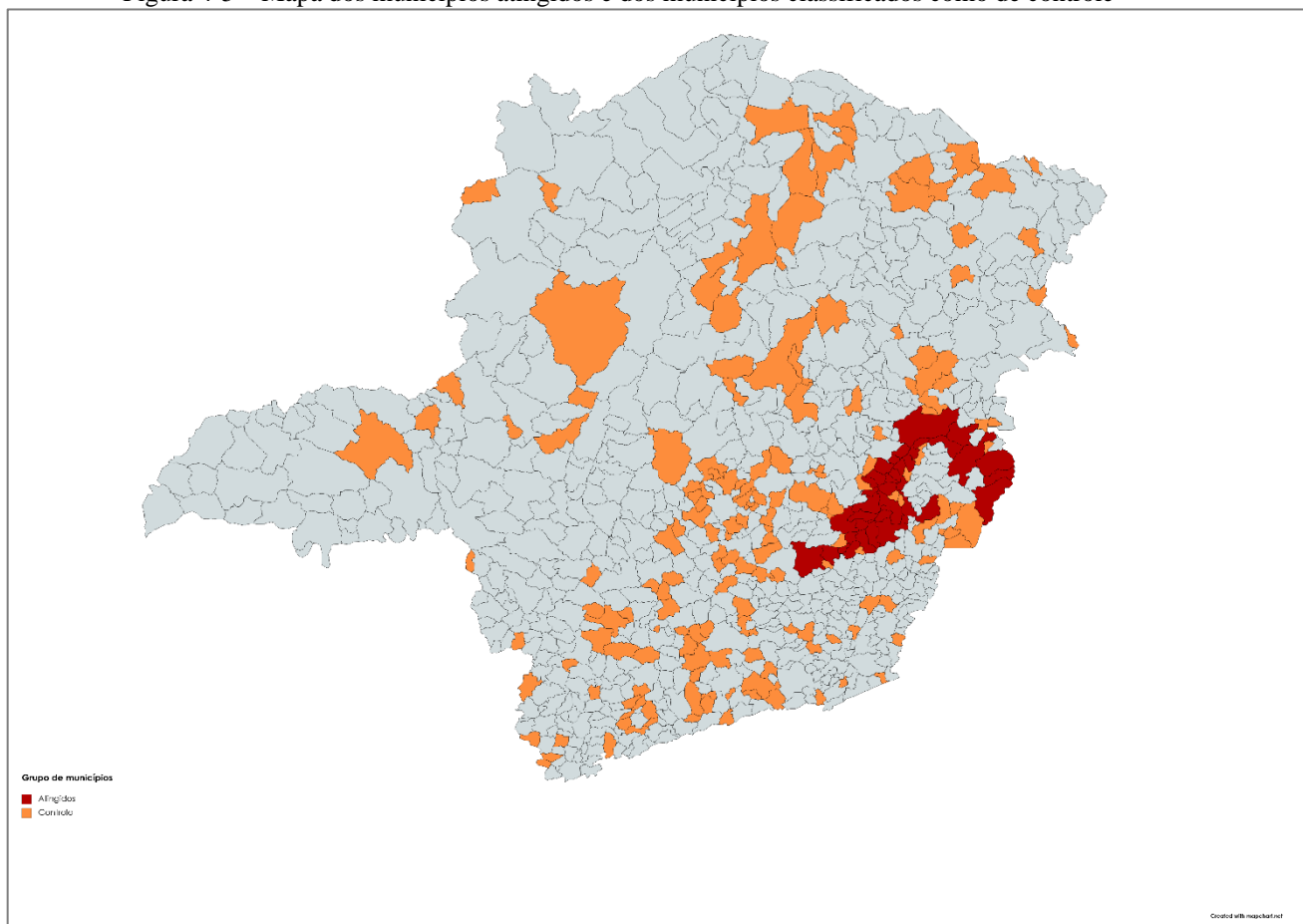
Ao recombinar as duas amostras, a sequência de rodadas permitiu encontrar uma relação de 169 municípios classificados para o grupo de controle desta avaliação. A lista de municípios de ambos os grupos pode ser conferida na Tabela 4-9 e no mapa da Figura 4-5.

Tabela 4-9 – Relação de municípios classificados para os grupos de atingidos e de controle

Grupo de atingidos		Grupo de controle			
Aimorés	Abadia dos Dourados	Claraval	Itabira	Olaria	São João do Manteninha
Alpercata	Acaiaca	Conceição da Barra de Minas	Itabirito	Ouro Branco	São João do Oriente
Barra Longa	Aguanil	Conceição das Pedras	Itaobim	Padre Paraíso	São João Evangelista
Belo Oriente	Águas Vermelhas	Congonhas	Itutinga	Pará de Minas	São Joaquim de Bicas
Bom Jesus do Galho	Aiuruoca	Contagem	Jaíba	Paraopeba	São José da Lapa
Bugre	Alagoa	Coronel Fabriciano	Janaúba	Passa Vinte	São José da Safira
Caratinga	Angelândia	Couto de Magalhães de Minas	Jequitaiá	Passabém	São José da Varginha
Conselheiro Pena	Arantina	Crucilândia	Jesuânia	Pedra Azul	São José do Mantimento
Córrego Novo	Augusto de Lima	Desterro de Entre Rios	João Pinheiro	Pedralva	São Sebastião do Rio Preto
Dionísio	Baldim	Desterro do Melo	Lagoa Dourada	Pedro Leopoldo	Senador Cortês
Fernandes Tourinho	Barão do Monte Alto	Diamantina	Lagoa Santa	Pequi	Seritinga
Galiléia	Barroso	Divinópolis	Lajinha	Piedade de Ponte Nova	Serra dos Aimorés
Governador Valadares	Bela Vista de Minas	Divisa Alegre	Lambari	Pimenta	Serrania
Iapu	Belo Vale	Dom Silvério	Lima Duarte	Piraúba	Serro
Ipaba	Betim	Dom Viçoso	Madre de Deus de Minas	Poços de Caldas	Sete Lagoas
Ipatinga	Bias Fortes	Engenheiro Caldas	Manhumirim	Pompéu	Simonésia
Itueta	Boa Esperança	Entre Folhas	Maravilhas	Porteirinha	Tabuleiro
Mariana	Bom Repouso	Ervália	Maria da Fé	Prados	Taiobeiras
Marliéria	Bonfim	Estrela D'Alva	Mata Verde	Prudente de Moraes	Taparuba
Naque	Brazópolis	Estrela do Sul	Mateus Leme	Raposos	Taquaraçu de Minas
Periquito	Cabeceira Grande	Felisburgo	Mathias Lobato	Ribeirão Vermelho	Toledo
Pingo D'Água	Caetanópolis	Fortuna de Minas	Matipó	Rio Acima	Três Pontas
Raul Soares	Camacho	Francisco Dumont	Mato Verde	Rio Manso	Turvolândia
Resplendor	Cambuquira	Francisco Sá	Matozinhos	Sabará	Uberlândia
Rio Casca	Campanário	Frei Inocêncio	Mercês	Sabinópolis	Umburatiba
Rio Doce	Campos Gerais	Glaucilândia	Mesquita	Salinas	Uruana de Minas
Santa Cruz do Escalvado	Carbonita	Goiabeira	Minduri	Santa Bárbara do Monte Verde	Vargem Alegre
Santana do paraíso	Carmo da Cachoeira	Gonzaga	Miradouro	Santa Cruz de Salinas	Varjão de Minas
São Domingos do Prata	Carmo da Mata	Guaxupé	Monte Azul	Santana do Deserto	Vespasiano
São José do Goiabal	Carmo de Minas	Guidoval	Monte Sião	Santana do Riacho	
São Pedro dos Ferros	Carmo do Paranaíba	Guimarânia	Montes Claros	Santo Antônio do Amparo	
Sem-Peixe	Carrancas	Ibertioga	Munhoz	Santo Hipólito	
Sobralia	Carvalhos	Ibituruna	Mutum	São Francisco de Paula	
Timóteo	Central de Minas	Ilicínea	Nazareno	São Gonçalo do Pará	
Tumiritinga	Chalé	Ipanema	Nova Era	São João da Lagoa	

Fontes: Censo 2010 e IMRS 2014.

Figura 4-5 – Mapa dos municípios atingidos e dos municípios classificados como de controle



Criado pelo mapchart.net²².

²²Este mapa foi licenciado pela Creative Commons Atribuição – CompartilhaIgual 4.0 Internacional (CC BY-SA 4.0).

5 ANÁLISE DOS RESULTADOS

Neste capítulo são analisados os principais resultados adquiridos na avaliação de impacto.

Na primeira parte, é elaborada uma descrição dos indicadores dos dois grupos ao longo do tempo. Dado o pressuposto do método de diferenças em diferenças, observar resultados inicialmente díspares não implica em disparidade entre as variações constatadas nos grupos. Por outro lado, estes dados não serão suficientes para avaliar impactos, sem o auxílio das próximas informações obtidas.

A segunda parte foca nos resultados do IVF globais por horizonte de tempo. Esta avaliação mostrou quais efeitos podem ter ocorrido na população dos municípios atingidos em condição de pobreza segundo os critérios do Cadastro Único.

A terceira parte constitui na avaliação decompondo as variáveis demográficas, formando os subgrupos e comparando suas variações nos horizontes de tempo especificamente nas dimensões de acesso ao trabalho e escassez de renda.

Por fim, haverá uma análise focada exclusivamente em como as diferentes composições familiares absorveram o impacto de longo prazo (2014 a 2018) do desastre.

5.1 Descrição dos municípios dos grupos de atingidos e de controle

Na Tabela 5-1 tem-se o tamanho das populações cadastradas no CadÚnico, para o grupo de atingidos, para o grupo de controle e para o estado de Minas Gerais como um todo. Dada a quantidade de municípios classificados como controle, é natural que seu contingente de famílias envolvidas seja maior.

Percebe-se que todos passam por um movimento de tendências oscilantes ao longo da década, mas com os três grupos seguindo os mesmos movimentos. Na Figura 5-1 podem-se observar as taxas de crescimento anuais de cada grupo: aumento em mais de 5% dos cadastrados de 2012 a 2013 e de 2013-14, seguida de uma queda de cerca de 5% negativo em 2014-15, seguida de uma nova queda em 2016. Até 2015, as taxas estavam com comportamentos relativamente bem próximos. Contudo, a partir de então, o grupo de municípios atingidos tornou-se dissonante, ainda que respeitando as tendências. Entre 2015-16, enquanto o estado e o grupo de controle observaram desaceleração na queda nos cadastrados (-1,9% e -2,6%, respectivamente), os atingidos mantiveram o ritmo de redução de 2014-15. Já entre 2016-17,

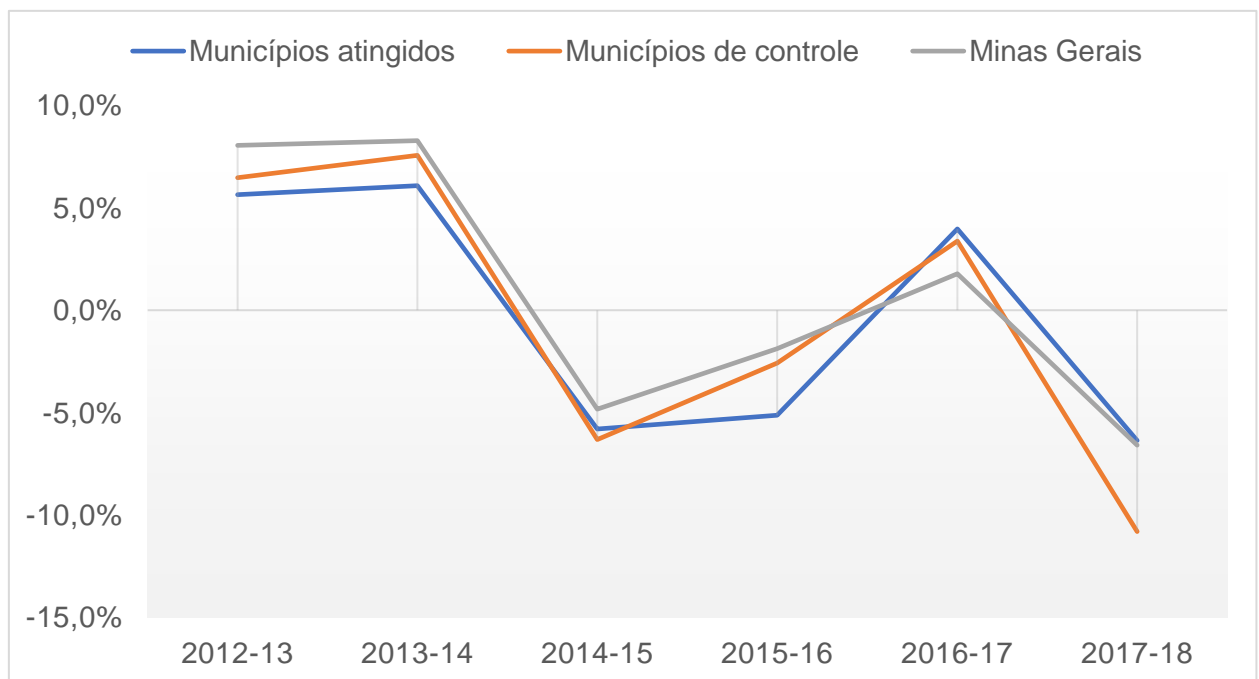
ocorre um novo aumento nas populações cadastradas dos três grupos, mas com o grupo de atingidos agora superando os demais: no estado todo fora 1,8%, no grupo de controle 3,4% e no grupo de atingidos 4%. Já na última variação avaliada, 2017-18, os atingidos diminuem com uma taxa igual à do estado (-6,5%), enquanto o grupo de controle tem a maior queda do período, de -10,8%.

Tabela 5-1 – N° de famílias cadastradas no CadÚnico entre 2012 e 2018, por grupo

N° de famílias	2012	2013	2014	2015	2016	2017	2018
Municípios atingidos	28.397	29.999	31.881	30.085	28.582	29.738	27.904
Municípios de controle	120.206	128.246	138.326	129.864	126.557	130.901	117.485
Minas Gerais	4.582.171	4.966.694	5.395.665	5.141.037	5.045.073	5.135.674	4.807.996

Fonte: CadÚnico (base desidentificada).

Figura 5-1 – Taxas de crescimento anuais da população cadastrada no CadÚnico, entre 2012 e 2018, por grupo



Fonte: CadÚnico (base desidentificada).

Segue na Tabela 5-2 os resultados médios dos municípios de Minas Gerais, dos 35 municípios atingidos, e dos 116 municípios de controle, nos valores utilizados como parâmetros para classificar o grupo de controle. Esta é uma medida um pouco diferente da avaliação per si: enquanto aqui se observa o resultado de cada município isoladamente e se somando em média, a avaliação será, na verdade, utilizando a totalidade da população cadastrada no CadÚnico de todos os municípios do grupo somados, como se fosse uma única unidade territorial.

Tabela 5-2 – Resultados das variáveis de parâmetro para os grupos de municípios atingidos, de controle e do estado de Minas Gerais

Variáveis	MG	Atingidos	Controle
Taxa de urbanização	70	76	78
Taxa de atendimento da educação básica	90	89	95
Proporção da população atendida pela ESG	91	92	90
IGDSUAS	83	94	78
V1. Alguma mulher teve filho nascido vivo no último ano	0,04 5	0,043	0,047
V2. Alguma mulher teve filho nascido vivo nos últimos dois anos.	0,14 5	0,142	0,148
V3. Presença de criança.	0,40 4	0,401	0,411
V4. Presença de criança ou adolescente.	0,52 0	0,518	0,529
V5. Presença de criança, adolescente ou jovem.	0,59 2	0,588	0,603
V6. Presença de idoso	0,23 4	0,250	0,221
V7. Ausência de cônjuge.	0,20 1	0,205	0,213
V8. Menos da metade dos membros encontram-se em idade ativa	0,02 6	0,024	0,033
V9. Presença de pessoas com dificuldade grande ou impossibilidade de caminhar, enxergar, ouvir ou com deficiência mental permanente.	0,19 6	0,219	0,211
V10. Presença criança no domicílio que não viva com a mãe	0,01 0	0,009	0,010
T1. Menos da metade dos membros em idade ativa encontram-se ocupados	0,02 6	0,038	0,033
T3. Ausência de ocupado no setor formal	0,59 0	0,587	0,574
T4. Ausência de ocupado em atividade não agrícola	0,45 6	0,454	0,430
T5. Ausência de ocupado com rendimento superior a um salário-mínimo	0,58 7	0,614	0,591
T6. Ausência de ocupado com rendimento superior a dois salários-mínimos	0,84 0	0,853	0,84
R1. Renda familiar per capita inferior à linha de extrema pobreza	0,12 8	0,125	0,126
R2. Renda familiar per capita inferior à linha de pobreza	0,30 9	0,326	0,316
R3. Maior parte da renda familiar advém de transferências	0,04 2	0,039	0,036

Fonte: Censo 2010; IMRS 2014.

Analisando especificamente o Índice de Vulnerabilidade das Famílias e suas dimensões para os grupos na Tabela 5-3, percebe-se que os municípios de ambos os grupos possuem IVF muito próximos dentre si: por exemplo, em 2012 temos 0,415 para os municípios atingidos e 0,416 para o grupo de controle. Além disso, é notável que se trata de grupos cuja vulnerabilidade sempre foi menor que a total do estado de Minas Gerais.

Tabela 5-3 – Resultados anuais do IVF, por grupos de atingidos e de controle; entre 2012 a 2018

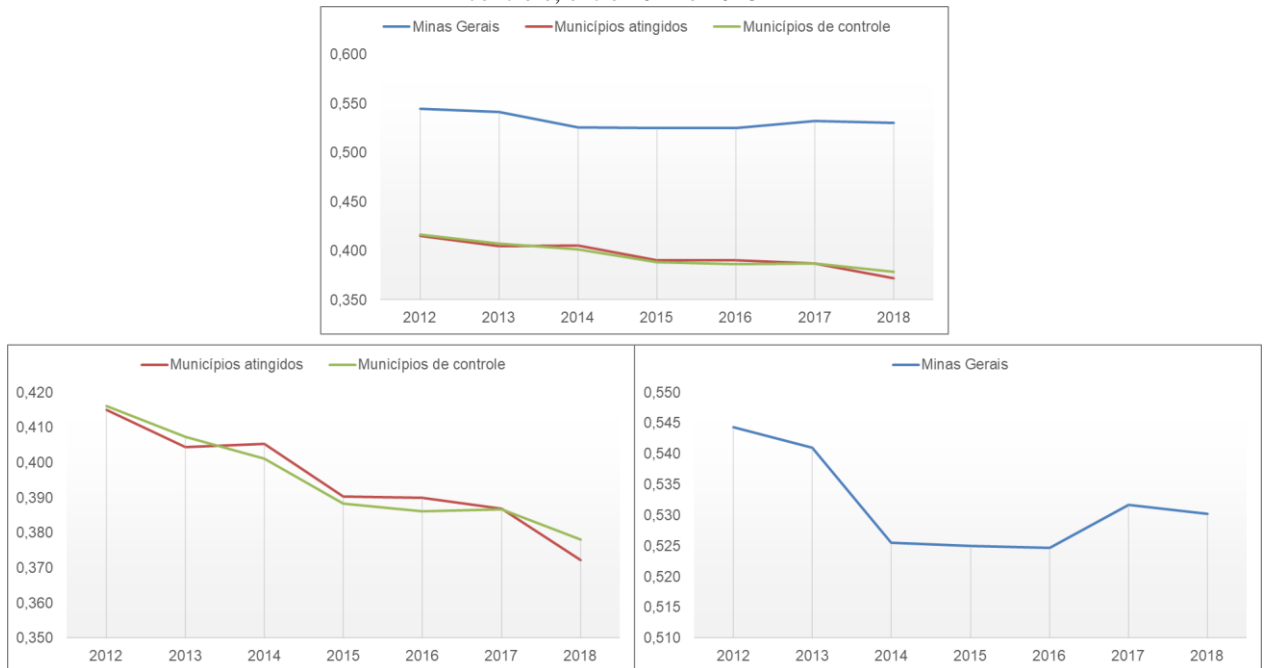
Minas Gerais		2012	2013	2014	2015	2016	2017	2018
IVF	Índice global	0,421	0,414	0,405	0,397	0,397	0,395	0,386
T	Acesso ao trabalho	0,544	0,541	0,526	0,525	0,525	0,532	0,530
R	Escassez de recursos	0,756	0,739	0,727	0,706	0,715	0,708	0,683
C	Acesso ao conhecimento	0,675	0,662	0,641	0,627	0,614	0,606	0,602
D	Desenvolvimento infantojuvenil	0,159	0,162	0,170	0,167	0,169	0,169	0,123
H	Condições habitacionais	0,160	0,156	0,143	0,138	0,136	0,130	0,123
V	Vulnerabilidade social	0,232	0,223	0,227	0,218	0,220	0,226	0,255
Municípios atingidos		2012	2013	2014	2015	2016	2017	2018
IVF	Índice global	0,415	0,404	0,405	0,390	0,390	0,387	0,372
T	Acesso ao trabalho	0,523	0,521	0,526	0,507	0,511	0,516	0,512
R	Escassez de recursos	0,756	0,740	0,737	0,716	0,720	0,713	0,675
C	Acesso ao conhecimento	0,656	0,648	0,644	0,612	0,602	0,587	0,579
D	Desenvolvimento infantojuvenil	0,156	0,167	0,160	0,172	0,175	0,174	0,124
H	Condições habitacionais	0,157	0,120	0,136	0,110	0,106	0,104	0,093
V	Vulnerabilidade social	0,242	0,230	0,229	0,225	0,225	0,226	0,249
Municípios de controle		2012	2013	2014	2015	2016	2017	2018
IVF	Índice global	0,416	0,407	0,401	0,388	0,386	0,387	0,378
T	Acesso ao trabalho	0,530	0,528	0,519	0,510	0,507	0,517	0,514
R	Escassez de recursos	0,751	0,736	0,728	0,698	0,707	0,703	0,682
C	Acesso ao conhecimento	0,660	0,645	0,623	0,606	0,590	0,585	0,583
D	Desenvolvimento infantojuvenil	0,158	0,161	0,169	0,167	0,166	0,169	0,124
H	Condições habitacionais	0,161	0,148	0,140	0,128	0,123	0,118	0,109
V	Vulnerabilidade social	0,237	0,226	0,228	0,222	0,223	0,228	0,257

Fonte: CadÚnico (base desidentificada).

Ao analisar a evolução do indicador ao longo do período, tem-se também diferenças entre os grupos e o estado. Na Tabela 5-3, é possível identificar que a trajetória da vulnerabilidade no estado foi dec linante entre 2012 e 2014, estagnando até 2016, e então regredindo para uma maior vulnerabilidade em 2017 e 2018. No entanto, os dois grupos avaliados, o movimento de queda da vulnerabilidade foi bem similar dentre ambos, e com uma tendência de queda mais forte e estável.

Percebe-se que os municípios atingidos possuíam uma vulnerabilidade menor em 2012 e 2013. Mas em 2014, um ano antes do rompimento da barragem de Fundão, estes municípios já apresentaram uma piora no indicador, passando a ser o grupo mais vulnerável até 2017, onde empataram no IVF e, no ano seguinte, o grupo de atingidos voltou a ser menos vulnerável.

Figura 5-2 – Índice de Vulnerabilidade das Famílias para o estado de Minas Gerais e grupos de atingidos e de controle, entre 2012 e 2018



Fonte: CadÚnico (base desidentificada).

Contudo, como observado anteriormente, esta análise descritiva não é suficiente para se constatar evidências de impactos. A partir do próximo tópico, serão abordadas as análises em que se faz a comparação das variações destes indicadores, bem como suas dimensões e algumas de suas componentes.

5.2 Impactos no Índice de Vulnerabilidade das Famílias

A análise abaixo se refere aos resultados da comparação entre toda a população registrada no Cadastro Único, por ano, para os dois grupos de municípios: atingidos e de controle.

Além do índice global e de suas seis dimensões principais, avalia-se também os efeitos especificamente nas variáveis componentes das dimensões de acesso ao trabalho e escassez de renda.

5.2.1 Diferenças em diferenças por período

A avaliação do impacto nos municípios atingidos sugere efeitos negativos para a vulnerabilidade das famílias nos prazos mais curtos, ou seja, o IFV se deteriora nos períodos 2014-15 e 2014-16, que são superados e invertidos para efeitos positivos quando se comparam

variações nos prazos mais longos (2014-17 e 2014-18). Ou seja, há uma tendência que aparenta um forte choque inicial, mas que, com o passar dos anos, a vulnerabilidade das famílias não só diminuiu como conseguiu progredir melhor do que nos municípios de controle. Ressalta-se, no entanto, que as diferenças em diferenças apresentam distâncias muito pequenas para o indicador agregado, ou seja, é preciso cautela para qualificar esta evidência de impacto como algo relevante. Neste sentido, a análise desagregada se torna ainda mais necessária.

Analisando por dimensões, têm-se impactos positivos em várias delas: desenvolvimento infantojuvenil; condições habitacionais; vulnerabilidade social; e escassez de recursos. O que quer que tenha acontecido com a população agregada do CadÚnico nos municípios atingidos, tornou-os mais resistente ao desastre e talvez as políticas de mitigação possam ter algum efeito que tenha impulsionado esse movimento.

Por outro lado, as dimensões que demonstraram impacto negativo, acesso ao trabalho e acesso ao conhecimento por adultos, foram persistentes nos primeiros anos. O acesso ao trabalho sofreu efeitos negativos até três anos após o desastre. No longo prazo, no entanto, ele também progride e passa a ter efeitos minimamente positivos em relação aos municípios de controle. O acesso ao conhecimento dos adultos, que envolve abandono escolar e analfabetismo, persiste negativo em todos os períodos avaliados.

Ou seja, há evidência de impacto do desastre na vulnerabilidade das famílias, negativo no curto prazo, mas que foi se recuperando ao longo dos anos, distribuído de formas diferentes dentre as seis dimensões do IVF.

5.2.2 Decomposição das dimensões de acesso ao trabalho e escassez de renda

Sendo o foco deste estudo as relações entre ocupação e pobreza, é interessante aprofundar a análise especificamente nas dimensões de acesso ao trabalho e de escassez de renda. Mas avaliar cada variável isoladamente pode demonstrar elasticidades muito distintas.

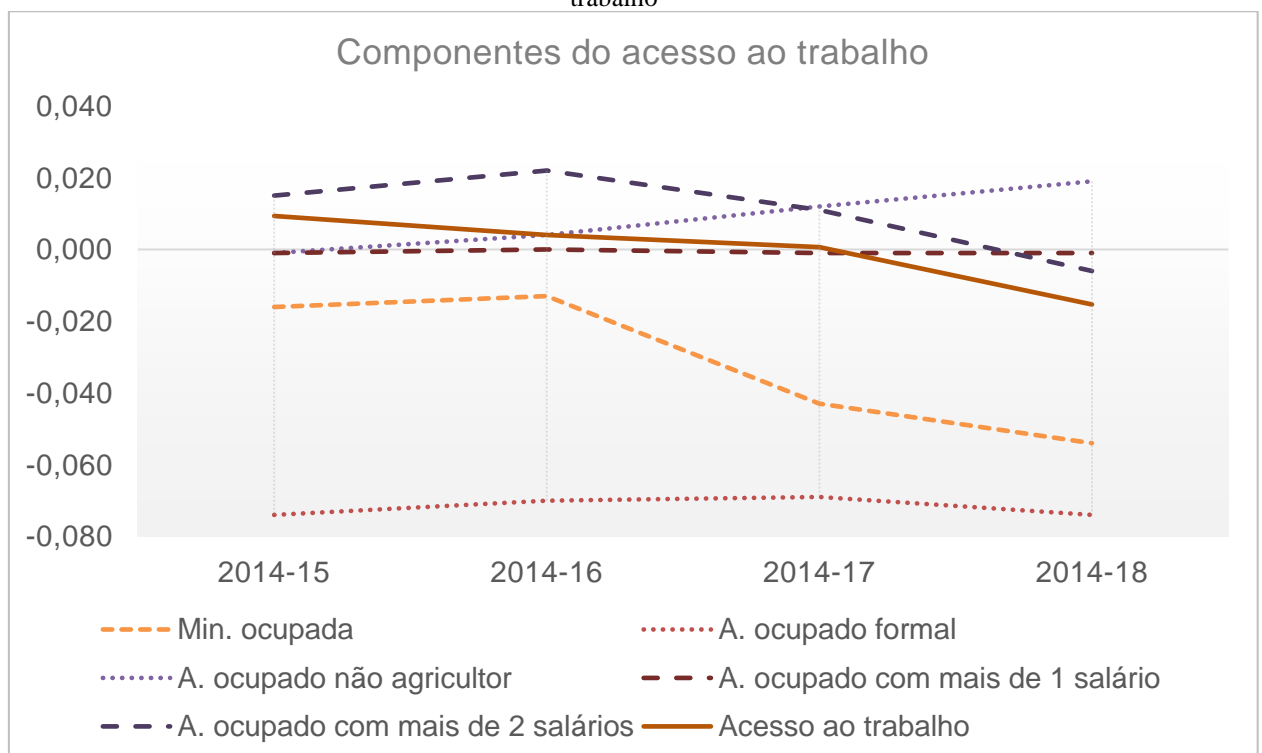
Enquanto o IVF e suas dimensões são agregados de valores médios, cada componente consiste em uma variável binária, e seus critérios podem afetar o número de observações considerado: algumas analisam a presença ou ausência de um ou mais indivíduos, enquanto outras somam valores de todos eles antes de atribuir a condição de vulnerabilidade. Por exemplo, a presença de “ao menos uma pessoa recebendo um salário-mínimo” depende da observação de um indivíduo no domicílio que esteja ocupado e o quanto ele, sozinho, recebe de renda. A “prevalência da fonte de renda do trabalho sobre as demais”, por seu turno, é calculada a partir da soma dos rendimentos de todos os membros do domicílio, trabalhando ou

não. Além disso, pela forma como o IVF foi estruturado para medir riscos, em alguns casos elas são acumuladas em cascata (veja em 3.6). Por exemplo, a variável de condição de extrema pobreza está contida na variável de condição de pobreza. Portanto, algumas análises precisam de muito cuidado para se identificar efeitos altamente correlacionados.

As ações de mitigação do desastre, tanto pelo poder público quanto pela própria Fundação Renova, podem ter efeitos diversos em variáveis mais específicas, dependendo de seu público-alvo e da extensão de sua execução. Sem uma avaliação de impacto específica para estes programas, torna-se impraticável isolá-los da análise: pressupõe-se que as ações de mitigação também são uma consequência do próprio desastre. Portanto, a análise por variáveis será voltada mais para se gerar novas hipóteses, que poderiam ser analisadas em melhor profundidade com pesquisas direcionadas especificamente para cada uma delas.

5.2.2.1 Acesso ao trabalho

Figura 5-3 – Impacto de diferenças em diferenças, por período, para as componentes da dimensão acesso ao trabalho



Fonte: CadÚnico.

A presença de uma maioria dos membros em idade ativa ocupados progrediu ao longo do período, mesmo no curto prazo (Figura 5-3). A partir do médio para o longo prazo, os domicílios no agregado conseguiram sustentar bem seus empregos disponíveis, por exemplo,

mantendo uma estabilidade no impacto positivo para a presença de ocupados com empregos formais. Contudo, um efeito mais aparente está na ausência de ocupados em setores não-agrícolas, que se agrava no médio e longo prazo, o que pode indicar, em algum grau, que os domicílios alteraram a estrutura de suas ocupações.

A presença de ocupados com um salário-mínimo ou mais não sofreu impactos relevantes, mas para os domicílios com alguém recebendo dois salários houve efeitos negativos. Ou seja, a faixa da presença de pessoas recebendo entre um e dois salários-mínimos é que parece ter se sustentado, enquanto domicílios com ocupados em melhores empregos precisaram receber rendimentos em menores patamares. Este cálculo usa como referência o salário-mínimo de referência para cada ano deflacionado pelo IPCA.

Por fim, um efeito que pode ocorrer é que, para se calcular a ausência de ocupados formais, utilizaram-se apenas domicílios onde havia algum ocupado, para então qualificar esta ocupação. Nesse sentido, se houve uma redução no número de ocupados informais que não tenha sido acompanhada por uma redução dos formais, é possível que a média esconda uma condição de efeitos desiguais, em que os trabalhadores informais possam ter sofrido mais que os formais. Surge então, hipóteses de interesse a serem abordadas, como por exemplo uma deterioração dos empregos oferecidos, como uma sustentação de pessoas em empregos formais, mas com salários provavelmente menores, e uma maior ausência de ocupados em setores não-agrícolas, o que aumenta a vulnerabilidade destes domicílios.

5.2.2.2 *Escassez de recursos*

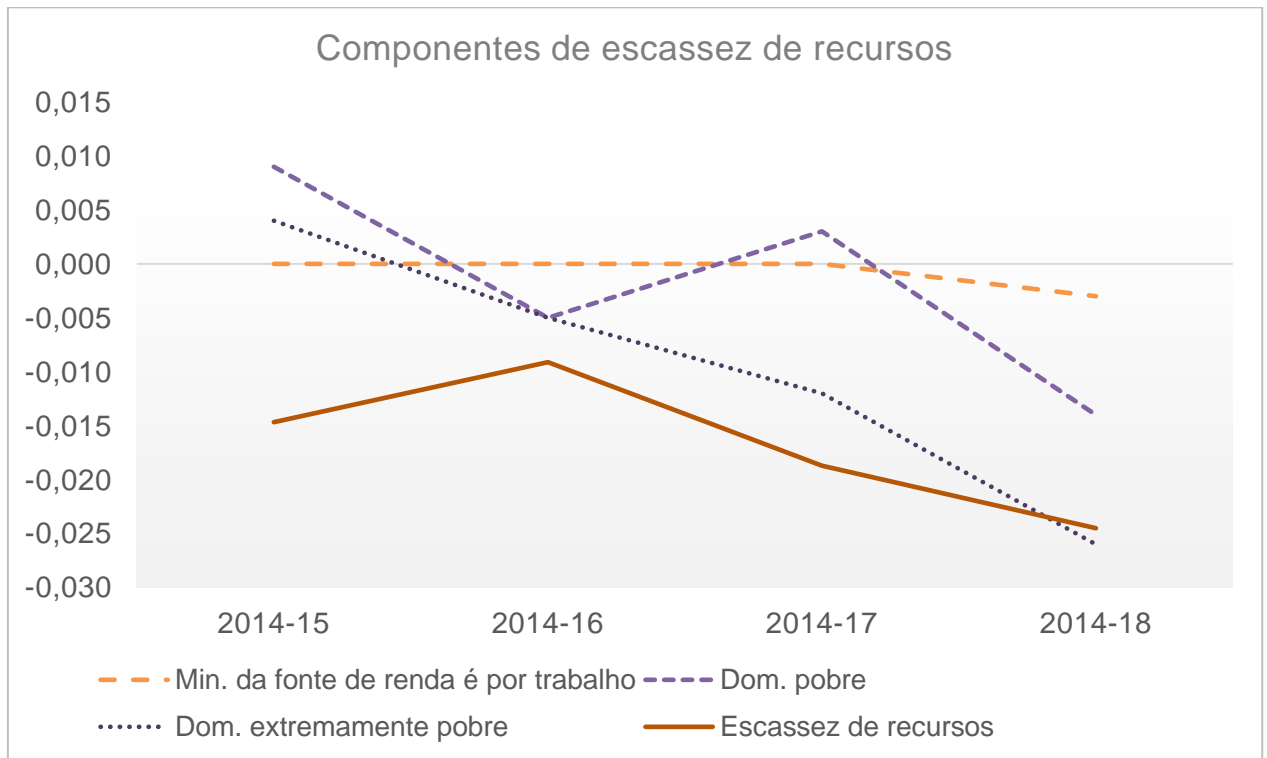
Há impactos positivos, quando se avalia em níveis agregados, nas componentes de escassez de renda. Uma redução na extrema pobreza pode ser observada, após um primeiro impacto negativo. A condição de pobreza oscilou, com um impacto negativo nos períodos de 2014-15 e 2014-17, mas que não se sustentou no longo prazo. O rendimento pelo trabalho ser minoritário dentre as fontes de renda não sofreu impactos relevantes.

É interessante notar um efeito na condição de pobreza que segue uma tendência parecida com a de extrema pobreza, exceto em um período (2014-16), em que há um ponto de inflexão com um impacto expressivo na presença de pessoas vivendo em domicílios considerados pobres. É um efeito parecido com o caso dos salários-mínimos: as faixas maiores sofreram mais no agregado que as faixas menores, tanto de pobreza quanto de renda do trabalho.

Como a base de dados utilizada foi o Cadastro Único, não é possível atestar que esse aumento de condição de pobreza seja generalizado para toda a população, mas seria interessante

verificar esta hipótese em especial, porque um dos requisitos de entrada de famílias no Cadastro Único é exatamente a condição de pobreza relacionada ao salário-mínimo nominal.

Figura 5-4 – Impacto de diferenças em diferenças, por período, para as componentes da dimensão escassez de renda



Fonte: CadÚnico.

5.3 Impactos da dinâmica demográfica nas dimensões de trabalho e de renda

Uma vez avaliadas as diferenças em diferenças potencialmente causadas pelo rompimento da barragem no IVF e em suas componentes para toda a população registrada no Cadastro Único, faz-se abaixo uma análise mais específica, buscando isolar as famílias cujas componentes demográficas estejam em condição de vulnerabilidade (variável igual a 1) ou de não vulnerabilidade (variável igual a 0), conforme descrito com mais detalhes no item 4.3.3. Ou seja, para quatro variáveis de vulnerabilidade demográfica avalia-se o impacto do desastre dentre as famílias que viviam em tal condição de vulnerabilidade e entre as famílias que não viviam nesta mesma condição de vulnerabilidade. São elas:

- (a) Presença de filhos nascidos vivos no ano (V1);
- (b) Presença de idosos com mais de 65 anos (V6);
- (c) Ausência do cônjuge do responsável pela família (V7); e
- (d) Minoria do domicílio em idade ativa (V8).

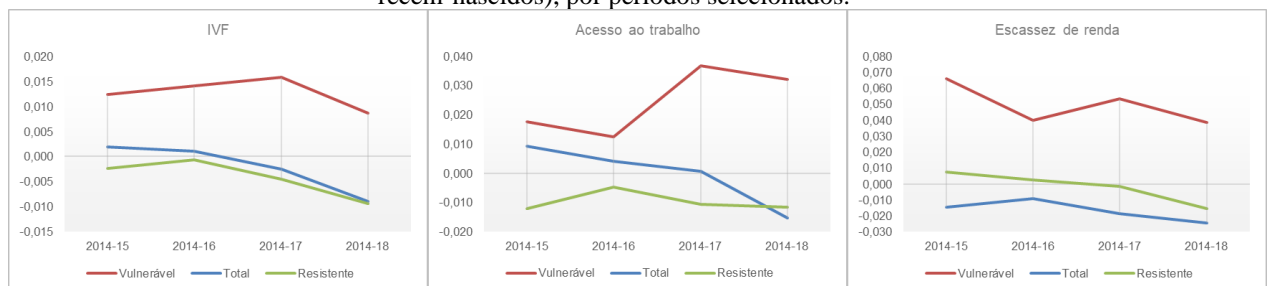
As famílias que não possuíam filhos ou idosos dentre seus membros; cujos responsáveis eram acompanhados por cônjuges; ou cuja maioria de seus membros estava em idade ativa, foram capazes de mitigar os efeitos negativos do desastre. Por outro lado, as famílias que de fato possuíam filhos recém-nascidos ou idosos; que não havia um cônjuge para o responsável; ou que a minoria dos membros estava em idade ativa, foram mais impactadas?

Cada tópico abaixo avaliará um destes casos, considerando os potenciais efeitos mitigadores para o IVF global e para suas dimensões de acesso ao trabalho e escassez de renda. Os resultados estão organizados ao final do capítulo na **Erro! Fonte de referência não encontrada..**

5.3.1 Presença de filhos nascidos vivos no ano (V1)

A presença de nascidos vivos há menos de 12 meses claramente tem fortes impactos no IVF, gerando diferenças de vulnerabilidade expressivas e persistentes ao longo do tempo (Figura 2-1).

Figura 5-5 – Impacto no grupo de atingidos por condição de vulnerabilidade pela variável V1 (presença de filhos recém-nascidos), por períodos selecionados.



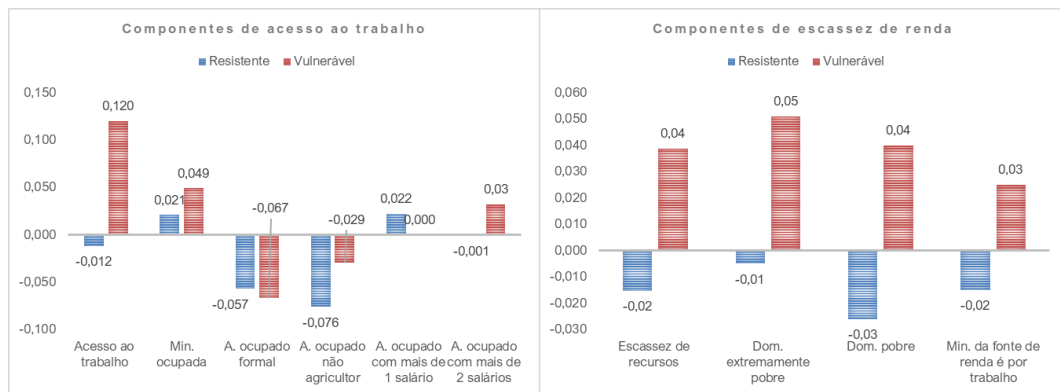
Fonte: CadÚnico. Elaboração própria

Quando se analisam as famílias sem recém-nascidos, percebe-se que elas demonstraram uma resistência maior no IVF, que só deixa de ser um impacto melhor que na população total em 2018. Por outro lado, as famílias com estas crianças sofreram impactos que se agravaram com o tempo, exceto no longo prazo (que ainda assim demonstra efeito negativo).

O acesso ao trabalho também demonstra que a presença de filhos com menos de um ano é um fator determinante das condições de vulnerabilidade do domicílio. Ao longo de todo o período, a vulnerabilidade por esta dimensão sofre impactos negativos, com um salto de aumento de vulnerabilidade após três anos do desastre (2014-17). No lado oposto, as famílias sem crianças de menos de um ano conseguiram efeitos positivos às condições de acesso ao trabalho, ao longo de todo o período.

Para o caso de escassez de renda, o efeito da presença de crianças é notável, causando um impacto maior, que arrefece ao longo do período. O subgrupo sem recém-nascido também sofreu impactos negativos, mas eles são revertidos no longo prazo. Quando se agregam os dois grupos, tem-se um efeito positivo no grupo de atingidos em relação ao grupo de controle²³.

Figura 5-6 – Determinantes das diferenças entre o impacto sofrido pelos subgrupos de resistentes e de vulneráveis pela variável V1 (presença de filhos recém-nascidos), no período entre 2014-2018



Fonte: CadÚnico.

Investigando as variáveis determinantes de cada dimensão, especificamente para o longo prazo (Figura 5-6), quando o domicílio possui uma criança nascida viva no último ano, surge um impacto negativo na vulnerabilidade de acesso ao trabalho que não existe para os não-vulneráveis. Os principais efeitos negativos foram nos casos de minoria dos membros do domicílio em idade ativa ocupados, e quanto à ausência de ocupados recebendo dois ou mais salários-mínimos (não houve impacto na condição de receber mais de um salário-mínimo).

Ambos os subgrupos observaram impactos positivos na ausência de ocupação formal, e este foi ainda melhor no grupo de vulneráveis. Ou seja, após o desastre, os municípios atingidos reduziram sua vulnerabilidade em relação ao trabalho formal, mesmo dentre os domicílios com a presença de crianças recém-nascidas. Já na condição de ausência de ocupados em setor não-agrícola, também houve impactos positivos, mas para os vulneráveis este foi menor.

Para as componentes da escassez de renda, no longo prazo, os vulneráveis por presença de filhos sofreram impactos negativos após o desastre, enquanto os não-vulneráveis tiveram

²³ Este tipo de efeito é possível porque os grupos possuem tamanhos diferentes, assim o impacto de um subgrupo não necessariamente “compensa” os efeitos sobre o outro. Como se trata de análise de diferenças em diferenças, a suscetibilidade pode ser diferente entre os subgrupos, ao ponto de que, na comparação global, a diferença da diferença se torna maior que nos subgrupos.

impactos positivos (ressalta-se que os níveis de impacto nesta dimensão são menores do que em relação ao acesso ao trabalho).

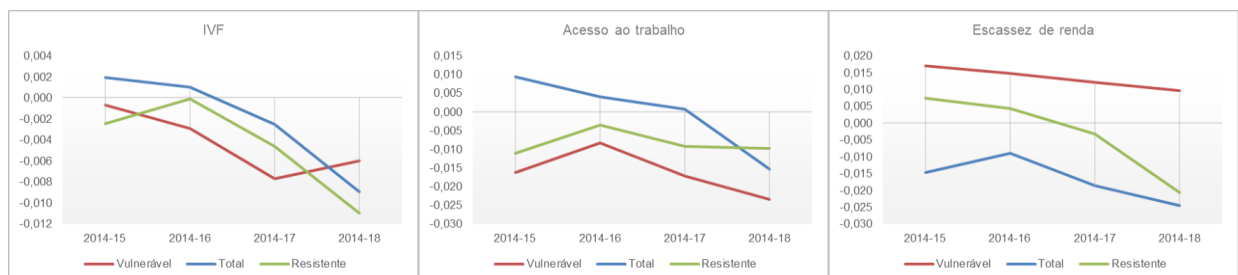
A condição de pobreza e de extrema pobreza foram agravadas de forma parecida. A condição de pobreza extrema dos vulneráveis em V1 foi mais impactada em relação ao seu subgrupo de controle (0,005); já a condição de pobreza em geral foi aquela que teve a maior disparidade com o resultado dos não-vulneráveis (0,007 de distância absoluta).

O mesmo aconteceu com a relação entre renda do trabalho e renda de outras fontes. Os não-vulneráveis em V1 obtiveram efeitos positivos no longo prazo, já aqueles domicílios com recém-nascidos receberam um impacto negativo de 0,03 em relação ao subgrupo de controle.

5.3.2 Presença de idosos com mais de 65 anos (V6)

Em geral, a presença ou não de idoso maior de 65 anos no domicílio demonstra impactos ambíguos no IVF ao longo do período. Em um primeiro momento, a presença de idoso agrava a condição de vulnerabilidade, enquanto os domicílios sem idosos mostraram inclusive algum efeito positivo. Este efeito se inverte quando se expande de 2014 até 2016. Três anos após o rompimento da barragem, a presença ou não de idoso deixa de causar impactos negativos.

Figura 5-7 – Impacto no grupo de atingidos por condição de vulnerabilidade pela variável V6 (presença de idosos com mais de 65 anos), por períodos seleccionados

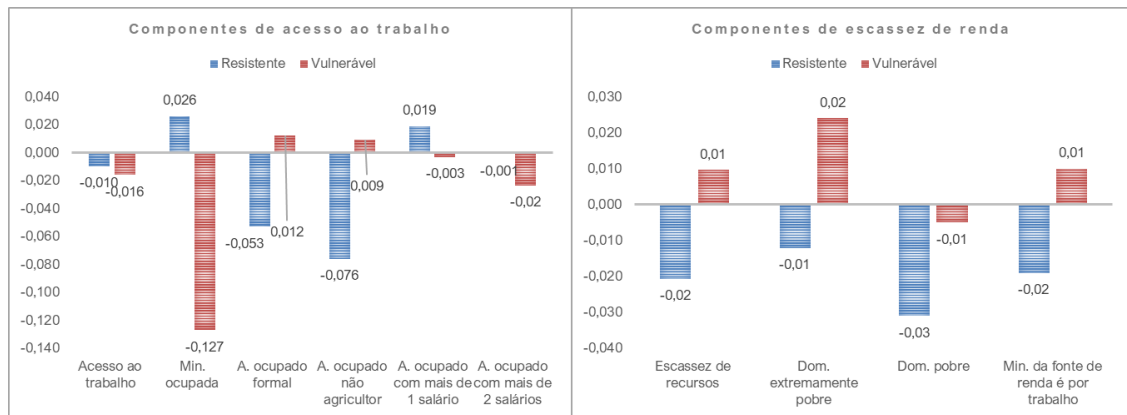


Fonte: CadÚnico.

Na dimensão de acesso ao trabalho, a presença de um idoso se mostra uma condição positiva após o desastre ao longo de todo o período, sendo interessante nota que, enquanto se compara o total dos cadastrados há um impacto negativo nos dois primeiros anos, nenhuma das comparações entre os subgrupos demonstra efeito similar.

Já na escassez de recursos, ter um idoso em casa de fato mostra um impacto negativo ao longo de todos os anos avaliados. Embora o subgrupo sem idosos também tenha sofrido impactos negativos, estes foram menores e chegam a ser positivos no longo prazo.

Figura 5-8 – Determinantes das diferenças entre o impacto sofrido pelo grupo total de atingidos e do subgrupo de resilientes pela variável V6 (presença de idosos com mais de 65 anos), no período entre 2014-2018



Fonte: CadÚnico. Elaboração própria

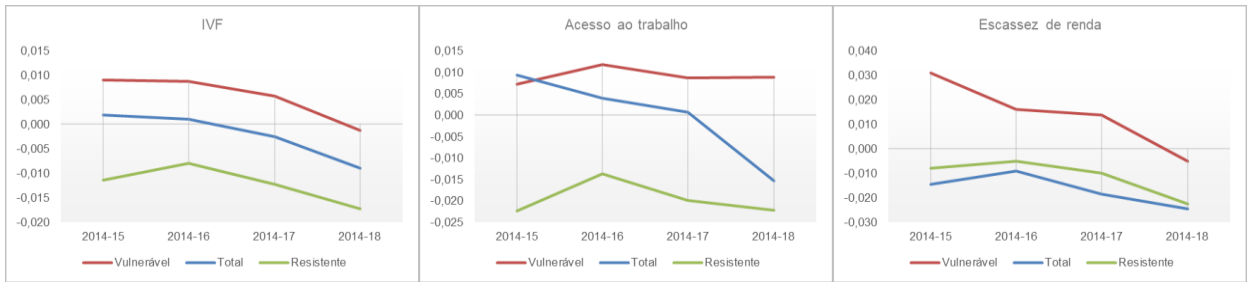
No longo prazo, é possível identificar que, com relação ao acesso ao trabalho, os grupos que possuíam idosos conseguiram impactos positivos em relação a manter a maioria dos membros em idade ativa no domicílio ocupados, melhor do que aqueles sem idosos. Em menor grau, o mesmo tipo de efeito parece ter acontecido com relação à presença de ocupados recebendo pelo menos um ou dois salários-mínimos, embora, com alguns efeitos irrelevantes, como pode ser visto na

Figura 5-8. O inverso ocorre em relação às variáveis de qualidade do emprego: as ausências de empregados formais ou de trabalhadores em setor não-agrícola são agravadas em domicílios com idosos, e progridem positivamente naqueles sem pessoas nesta faixa etária.

Com relação a escassez de recursos os efeitos são mais claros. A presença de idoso agrava a vulnerabilidade no longo prazo nas três variáveis, enquanto a ausência dele permite resultados melhores. Este padrão de impactos permite conjecturar a hipótese de que o idoso maior de 65 anos é, na prática, um indivíduo ainda capaz de trabalhar e que, na condição de aumento de risco de desemprego local, ele se torna uma mão de obra a mais do domicílio que poderia estar, até então, desocupada. Contudo, esta oferta de mão de obra não parece ter sido o suficiente para mitigar os riscos associados à perda de renda, tanto a entrada nas condições de pobreza quanto a renda advinda do trabalho, considerando que famílias com idosos aumentam suas chances de obterem renda através assistência social e previdência.

5.3.3 Ausência do cônjuge do responsável pela família (V7)

Figura 5-9 – Impacto no grupo de atingidos por condição de vulnerabilidade pela variável V7 (ausência do cônjuge do responsável pela família), por períodos selecionados



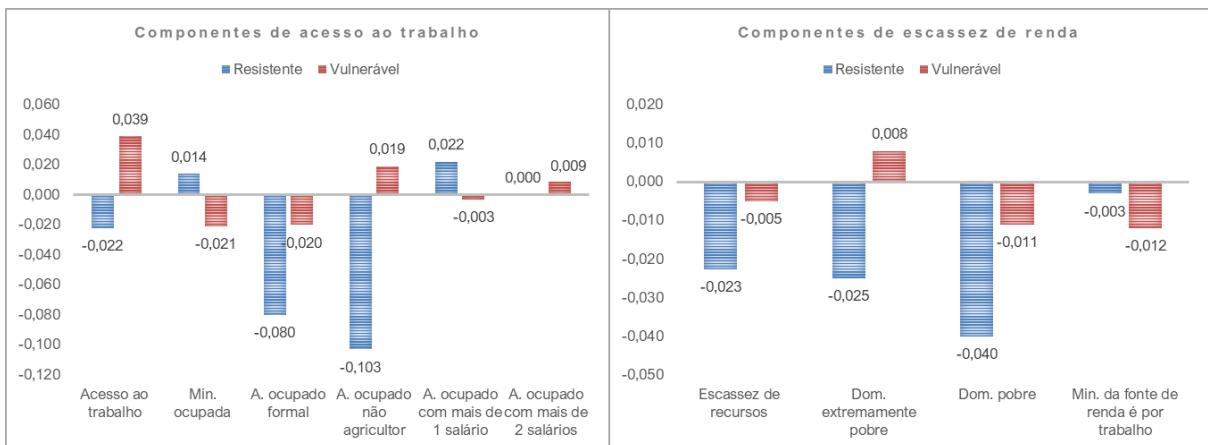
Fonte: CadÚnico.

As famílias em que não havia um cônjuge no domicílio foram as que mais aumentaram sua vulnerabilidade com o rompimento da barragem (Figura 5-9). O IVF total já demonstra a diferença, mas ela também se destaca nas vulnerabilidades pelo acesso ao trabalho e pela escassez de renda.

Quando há a presença de um cônjuge, temos uma demonstração de mitigação expressiva ao longo de todo o período. O IVF das famílias sem cônjuge demora dois anos a mais para tornarem os efeitos do desastre positivos, do que o todo dos cadastrados no CadÚnico. Enquanto isso, o subgrupo com cônjuge sempre conseguiu efeitos positivos no indicador global.

Movimento parecido ocorreu com a escassez de renda. Neste caso, apenas os subgrupos sem a presença de cônjuges demonstraram impactos negativos, que perduraram até 3 anos após o desastre. Em relação ao acesso ao trabalho, o subgrupo sem cônjuge se manteve ao longo de todo o período sofrendo impactos negativos, enquanto aqueles com cônjuges demonstraram um fenômeno oposto.

Figura 5-10 – Determinantes das diferenças entre o impacto sofrido pelo grupo total de atingidos e do subgrupo de resilientes pela variável V7 (ausência do cônjuge do responsável pela família), no período entre 2014-2018



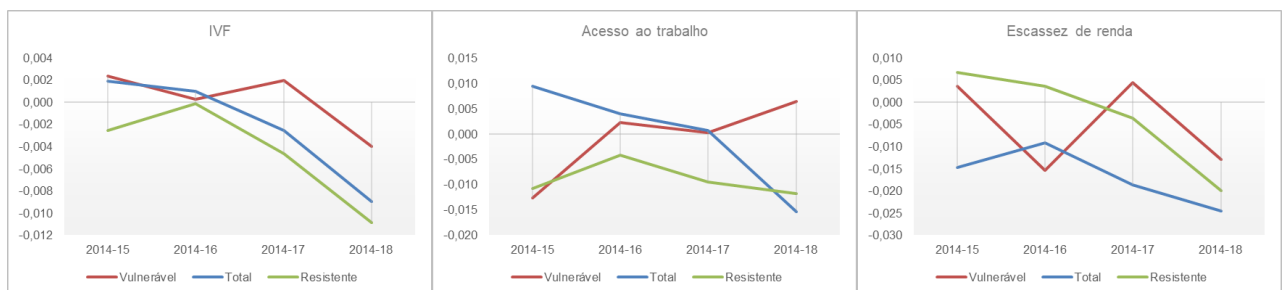
Fonte: CadÚnico. Elaboração própria

No longo prazo (Figura 5-10), a melhor resistência trazida pela presença do cônjuge foi no risco de ausência de ocupados em setores não-agrícolas ou formais. No entanto, nenhum impacto relevante parece ter sido observado em relação a salários, exceto para o grupo com presença de cônjuges, que tiveram um impacto negativo maior no grupo que recebe um ou mais salário-mínimo. A condição de minoria ocupada é agravada nos domicílios que possuíam cônjuges, e aliviada naqueles sem. Para a escassez de renda, os domicílios sem cônjuge sofreram mais a entrada na condição de pobreza, mas o inverso ocorreu com relação à condição de minoria da fonte de renda pelo trabalho.

Parece claro que famílias com casais conseguiram efeitos positivos do desastre, principalmente nas condições de renda e qualidade do emprego, mesmo que no geral tenham sofrido mais com o fato de que um dos dois tenha ficado sem emprego. O apoio do cônjuge parece ter sido fundamental para a manutenção e melhoria das condições de emprego e de pobreza dos domicílios.

5.3.4 Minoria do domicílio em idade ativa (V8)

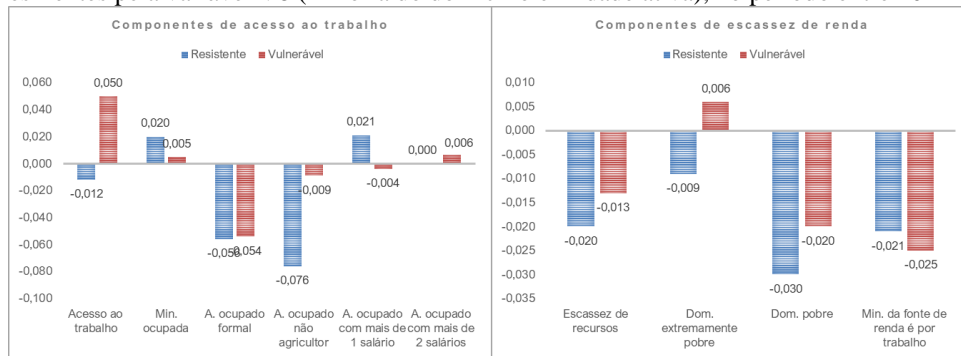
Figura 5-11 – Impacto no grupo de atingidos por condição de vulnerabilidade pela variável V8 (minoria do domicílio em idade ativa), por períodos selecionados



Fonte: CadÚnico.

Na Figura 5-11, observa-se que possuir a minoria dos membros do domicílio em idade ativa agravou o impacto do rompimento da barragem no IVF ao longo dos anos, efeito que só se encerra no longo prazo. Com relação ao acesso ao trabalho, é possível notar que possuir maioria em idade ativa permitiu impactos mais positivos e estáveis, enquanto aqueles minoritários tiveram impactos negativos na vulnerabilidade de acesso ao trabalho. Já na escassez de renda, os efeitos são ambíguos: o subgrupo com maioria sofre impactos negativos maiores nos dois primeiros anos, que se tornam positivos em 2014-17 e 2014-18, enquanto o subgrupo com minoria oscila ao longo do tempo.

Figura 5-12 – Determinantes das diferenças entre o impacto sofrido pelo grupo total de atingidos e do subgrupo de resilientes pela variável V8 (minorias do domicílio em idade ativa), no período entre 2014-2018



Fonte: CadÚnico.

No longo prazo (Figura 5-12), o acesso ao trabalho do subgrupo com maioria em idade ativa sofre mais impactos negativos quando se analisa se estas pessoas estão ocupadas. Com relação à ausência de ocupados formais, ambos os subgrupos tiveram o mesmo impacto positivo. Já para a ausência de ocupado em setor não-agrícola, aqueles com maioria em idade ativa sofreram impactos mais positivos do que os minoritários. A escassez de recursos mostra que, quando a maioria está em idade ativa, o impacto foi mais positivo em relação às condições de pobreza, mas menos positivo em relação à fonte de renda pelo trabalho.

6 CONSIDERAÇÕES FINAIS E CONCLUSÕES

Apesar do estudo de caso deste trabalho ser o rompimento da barragem de Fundão, seu objetivo principal foi mostrar o potencial que indicadores multidimensionais de vulnerabilidade social possuem em capturar os efeitos de um desastre tecnológico nas famílias atingidas, inclusive identificando diferenciais agravantes ou mitigadores na composição familiar ante mudanças súbitas e adversas nas condições de acesso ao trabalho e de escassez de recursos ao longo dos anos após o desastre.

Conforme descrito na análise bibliográfica, indicadores multidimensionais bem elaborados possuem atributos como a possibilidade de desagregação e de comparações de variáveis que originalmente possuem escalas diferentes. Além disso, permitem que analistas possam variar o critério de definição de linhas e de marcos de pobreza, de modo a focar suas análises em públicos-alvo bem definidos.

O uso do conceito de vulnerabilidade é importante, pois permite que a análise das condições sociais utilize a perspectiva de riscos ao invés de valores absolutos, ou seja, ajuda a capturar a evolução dos problemas sociais, independente da situação inicial. Assim, mesmo em uma situação inicialmente confortável que gradativamente se deteriora pode ser identificada, e efeitos mitigadores da composição familiar podem ser avaliados, que é uma informação muito relevante para a elaboração de políticas públicas de emergência e mitigação de prejuízos comunitários.

Este tipo de análise consegue mostrar como a demografia faz diferença em casos de desastres. Indicadores multidimensionais de vulnerabilidade que acrescentem componentes demográficas é capaz de avaliar o que a dinâmica familiar auxilia ou prejudica em situações de estresse. Elementos como a nupcialidade, a presença de mais membros em idade ativa, ou ainda a assistência social direcionada a dependentes são acionados nessas circunstâncias, permitindo aquelas pessoas que perderam fontes de trabalho e de renda encontrem suporte em outros membros do domicílio.

O Índice de Vulnerabilidade das Famílias, criado inicialmente por Barros, Carvalho e Franco (2006) e posteriormente aprimorado por Furtado (2012 e 2013), mostrou-se simples de se calcular e muito interessante para este estudo de caso, por ser inspirado diretamente pelas nossas pesquisas domiciliares e pelas características brasileiras. Sua adaptação para a base desidentificada do CadÚnico foi simples, com pouca necessidade de mudança, especialmente nas dimensões de acesso ao trabalho e escassez de recursos, o foco deste estudo. Contudo, elementos como condições de saúde ou de segurança ficaram de fora da avaliação, dada a falta

de informações consistentes na base. Na eventual análise com bases de dados mais sofisticadas para estas dimensões, talvez seja melhor a opção por novos índices; utilizando, por exemplo, o método de construção de indicadores Alkire-Foster (2011), que se tornou bastante conhecido e aplicado utilizado pela Pnud em seu Índice de Pobreza Multidimensional.

Especificamente para este estudo de caso, optou-se pela técnica de avaliação de impacto por diferenças em diferenças, comparando a variação do IVF calculado a partir dos dados do CadÚnico do grupo de municípios atingidos em relação à mesma variação para um grupo de municípios de controle. Os dados foram obtidos a partir de 2012 até 2018, observando o estado das famílias no dia 31 de dezembro de cada ano. Considerando o desastre ocorrido em novembro de 2015, assumiu-se como ponto de partida a variação entre dezembro de 2014 e dezembro de 2015, e em seguida as variações de períodos mais longos (2014-16, 2014-17 e 2014-18).

Para estabelecer um grupo de controle, foram utilizadas variáveis de parâmetro que identificaram um grupo de municípios mineiros semelhante ao grupo de municípios atingidos, de acordo com as dimensões do IVF de demografia, trabalho e renda, mais características estruturais dos municípios e de suas respectivas capacidades de atendimento à população. Usando o algoritmo *k-NN* ($k = 2$), classificou-se os 169 municípios mais próximos dos 35 considerados atingidos. Ao conferir a evolução do IVF destes dois grupos, percebeu-se que eles possuíam evoluções bem próximas um do outro, enquanto eram visivelmente diferentes do estado de Minas Gerais como um todo, o que se considerou um indício de que o grupo de controle foi razoavelmente bem definido. Mas é importante ressaltar que não houve um controle de significância estatística dos dados, ou seja, quaisquer evidências e resultados aqui descritos precisam ser confirmados em novos estudos antes de serem tratados como definitivos.

Este trabalho apresentou resultados aparentemente ambíguos, mas que conseguem traçar um cenário interessante para se analisar e debater. Considerando toda a população cadastrada anualmente no CadÚnico dos dois grupos de municípios, houve um aparente impacto negativo do desastre nos primeiros dois anos, que se inverteu nos anos seguintes. Mas estas variações foram todas na terceira casa decimal de um indicador que varia de 0 a 1, ou seja, é muito arriscado afirmar que isso seja um real impacto sendo observado. Da mesma forma, o acesso ao trabalho indica pequenos impactos negativos nos três primeiros anos, enquanto a escassez de renda mostra impactos positivos em todos os períodos.

Quando se isolam apenas as populações com certas características demográficas, tem-se diferenças marcantes. Dentre as famílias que possuem filhos recém-nascidos, houve aparentes impactos negativos por todos os períodos, tanto no IVF total quanto nas dimensões

de trabalho e de renda, com destaque negativo para a condição de pobreza ou de extrema pobreza do domicílio. Isto é uma possível evidência de que crianças limitam muito as condições das pessoas de se adaptar a mudanças abruptas, talvez dificultando buscas por novos empregos, ou exigindo demais da distribuição da renda no domicílio, dentre outras hipóteses.

Já as famílias com a presença de idosos parecem ter sofrido impactos positivos no IVF como um todo e na dimensão de trabalho, apesar de ainda sofrerem efeitos negativos com a escassez de renda. Isso pode acontecer porque, por um lado, pessoas com mais de 65 anos muitas vezes ainda são capazes de trabalhar; por outro, pode ser que elas não possuem o mesmo potencial de jovens para gerar uma renda em novos empregos que sustente o mesmo padrão de vida anterior.

A falta de cônjuge para o representante da família mostrou grande impacto negativo nos primeiros anos, mas que, para o IVF global e para a escassez de recursos, acaba por se inverter no longo prazo (2014-18). Já a dimensão de acesso ao trabalho é prejudicada por todos os períodos. É interessante levantar perguntas especificamente sobre o papel de cônjuges em casos de grande adversidade: é possível que sem um parceiro a pessoa não consiga buscar ou sustentar empregos com a mesma eficiência; talvez o fato de o IVF focar em condições domiciliares torna-o mais sensível à presença de mais adultos em idade ativa no domicílio; mudanças no comportamento do responsável quando se tem um parceiro; dentre outras possibilidades.

A última composição avaliada foi a presença de minoria em idade ativa do domicílio teve resultados mais voláteis. O IVF global pareceu demonstrar que as famílias com idade ativa minoritária foram mais impactadas negativamente, embora no longo prazo elas também se mostraram capazes de reverter a situação e tornarem este impacto algo positivo. No acesso ao trabalho, ainda no primeiro período (2014-15), as famílias com minoria ativa sofreram um provável impacto ligeiramente mais positivo. Contudo, este efeito se torna cada vez mais negativo com o tempo. O impacto da escassez de renda parece ser o mais volátil, mostrando que, dependendo do período, têm-se impactos negativos ou positivos para estas famílias, enquanto para as famílias com maioria em idade ativa tem-se um impacto negativo inicial, seguindo de progressos em cada período, tornando-se positivo a partir de 2014-17.

Mais especificamente, salta aos olhos um padrão: no longo prazo (2014-18), a variável que apontou para a presença de ocupados em empregos formais tendeu a sofrer impactos positivos com o desastre. No entanto, a ausência de ocupados recebendo mais de dois salários-mínimos, em valores reais de 2012 tendeu a se agravar (quando se considera presença de pessoas recebendo um ou mais salários mínimos, não houve impacto).

Este trabalho não tem a pretensão de identificar o que isto significa, sendo apenas possível especular hipóteses para próximos estudos. Contudo, é interessante notar paralelos com a literatura revista. Por um lado, os efeitos negativos do desastre da barragem de Fundão continua sendo extensamente investigados, mas resultados estatisticamente significativos precisam de dados de grande qualidade e análises muito precisas para que seja possível inferir as causas de eventuais impactos (REZENDE, 2021). Além disso, a demografia dos desastres leva tempo para ter seus efeitos no longo prazo, e o desastre foi analisado em um horizonte de quatro anos. Além disso, como pode se observar nos estudos revisados por Polizzi, Struffolino e Van Winkle (2016), um aspecto relevante nas relações entre demografia e trabalho em condição de pobreza é a entrada e saída das famílias nestas condições. O problema surge porque a informação obtida pela base de dados desidentificada do CadÚnico se mostra muito limitada, uma vez que se trata de um universo exclusivamente composto por famílias já consideradas pobres, e não é possível acompanhar trajetórias específicas para avaliar entradas e saídas do cadastro.

A continuidade do estudo da relação entre vulnerabilidade e demografia, especialmente diante de grandes adversidades, será muito fortalecida com o uso de bases de dados mais qualificadas. Por exemplo, a demografia dos desastres tem muito a nos informar caso use dados longitudinais para calcular indicadores multidimensionais de vulnerabilidade, pois conseguiria identificar a verdadeira trajetória dos impactos de um desastre nas famílias, descrevendo quais tipos de família correm mais riscos de entrar em condição de pobreza, e quais são aquelas que dão suporte bom o suficiente para mitigar efeitos negativos. Nesse sentido, o estudo de outros fenômenos, como resistência ou capital social, podem agregar muita informação sobre o que ocorre em cada dinâmica demográfica em situações adversas.

Outra limitação deste trabalho foi a incapacidade de separar os efeitos do rompimento da barragem propriamente dito das reações sociais e governamentais a ele. Por exemplo, ao se criar a Fundação Renova, mutirões de assistência, prioridade de destinação de recursos públicos, dentre outras respostas, é possível que tais reações tenham sido efetivas em atender a população atingida, ou tenha ido além e causado melhorias mais relevantes do que seria esperado caso não houvesse o desastre. Analistas de políticas públicas deveriam levar isto em conta, preferencialmente criando avaliações de impacto sobre as próprias ações implementadas.

Os analistas também podem aproveitar o conceito de vulnerabilidade social para uma mensuração mais apurada de elementos como riscos socioambientais de empreendimentos. É da natureza de técnicas de mensuração de risco que o resultado de seus indicadores seja aplicado como aumentos dos custos de investimento, e que suscite o planejamento de ações de mitigação:

risco de crédito, de mercado, ambiental, dentre outros, são parecidos com a estrutura dos indicadores multidimensionais de vulnerabilidade. Com o crescimento das práticas ambientais, sociais e de governança das organizações (ESG, da sigla em inglês para "*environmental, social and governance*"), este tipo de estudo talvez se torne mais comum nos próximos anos. Especificamente para políticas públicas de grande porte, como a construção de grandes obras de infraestrutura cujo planejamento projete sua execução por décadas, a inserção de variáveis demográficas no cálculo de vulnerabilidade pode se tornar um instrumento bastante valioso.

Por fim, é revelador que tenha ocorrido um segundo desastre ambiental de grandes proporções no estado em 2019, em Brumadinho, e que o mundo tenha passado por uma pandemia em 2020 que exigiu mudanças abruptas na vida de todos os brasileiros. Estes são outros dois fenômenos em que a vulnerabilidade social atrelada à demografia se mostrará determinante para as capacidades das pessoas de se sustentarem, para além da simples medida de pobreza por falta de dinheiro.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALAM, Bashir. **Implementing kNN algorithm using Python**. 2022. Hands-On-Cloud. Disponível em: <<https://hands-on.cloud/implementing-knn-algorithm-using-python/>>. Acesso em: 15 jul. 2022.
- ALKIRE, Sabina; FOSTER, James. Counting and multidimensional poverty measurement. **Journal Of Public Economics**, [Amsterdã, Países Baixos], v. 95, n. 7-8, p. 476-487, ago. 2011. Elsevier BV. Disponível em: <<http://dx.doi.org/10.1016/j.jpubeco.2010.11.006>>. Acesso em: 16 maio 2022.
- ALKIRE, Sabina; JAHAN, Selim. **Assessing multidimensional poverty: one index at the global level. one index at the global level**. 2018. Pnud. Disponível em: <<https://hdr.undp.org/content/assessing-multidimensional-poverty-one-index-global-level>>. Acesso em: 27 jul. 2022.
- ALVALÁ, Regina C. S.; BARBIERI, Alisson. Desastres naturais. In: NOBRE, Carlos A.; MARENGO, José A. (org.). **Mudanças climáticas em rede: um olhar interdisciplinar**. Bauru: Canal 6 Editora, 2017. Cap. 7. p. 203-230. Disponível em: <https://www.researchgate.net/publication/322592026_Mudancas_Climaticas_em_Rede_Um_olhar_interdisciplinar>. Acesso em: 26 jul. 2022.
- BAMBERGER, Michael; RAO, Vijayendra; WOOLCOCK, Michael. Using Mixed Methods in Monitoring and Evaluation: experiences from international development. **World Bank Policy Research Working Paper: No. 5245**. [S. I.], p. 1-30. 01 mar. 2010. Disponível em: <<https://ssrn.com/abstract=1578113>>. Acesso em: 18 mar. 2022.
- BANCO MUNDIAL. FAQs: global poverty line update. Global Poverty Line Update. 2015. Disponível em: <<https://www.worldbank.org/en/topic/poverty/brief/global-poverty-line-faq>>. Acesso em: 08 jul. 2022.
- BANCO MUNDIAL. **Quase Metade do Mundo Vive com Menos de USD \$5.50 por Dia**: apesar do declínio da pobreza extrema, macroindicadores mostram que milhões ainda lutam para satisfazer as necessidades básicas. Apesar do declínio da pobreza extrema, macroindicadores mostram que milhões ainda lutam para satisfazer as necessidades básicas. 2018. Disponível em: <<https://www.worldbank.org/pt/news/press-release/2018/10/17/nearly-half-the-world-lives-on-less-than-550-a-day-brazilian-portuguese.print>>. Acesso em: 07 jul. 2022.
- BARROS, Ricardo Paes de; CARVALHO, Mirela de; FRANCO, Samuel. **Pobreza Multidimensional no Brasil**. Rio de Janeiro, Rj: Ipea, 2006. 40 p. (Texto para discussão 1227). Disponível em: <https://www.ipea.gov.br/portal/index.php?option=com_content&view=article&id=4372>. Acesso em: 02 jul. 2022.
- BASE DOS DADOS. **Cadastro Único**. [2019]. Disponível em: <<https://basedosdados.org/dataset/cadastro-unico>>. Acesso em: 19 jul. 2022.
- BEENA, Vishwanath. **Understanding Confusion matrix and applying it on KNN-Classifer on Iris Data set**. 2020. Artificial Intelligence in Plain English. Disponível em:

<<https://ai.plainenglish.io/understanding-confusion-matrix-and-applying-it-on-knn-classifier-on-iris-dataset-b57f85d05cd8>>. Acesso em: 15 jul. 2022.

BIRD, Deanne; TAYLOR, Andrew. Demographic Change of ‘Single-Industry’ Towns: decline and resilience in Morwell, Australia. In: KARÁCSONYI, Dávid; TAYLOR, Andrew; BIRD, Deanne (ed.). **The Demography of Disasters**: impacts for population and place. [Nova Iorque, Eua]: Springer, 2021. Cap. 7. p. 125-151. Disponível em: <https://doi.org/10.1007/978-3-030-49920-4_9>. Acesso em: 26 jul. 2022.

BISSOLI JUNIOR, Alexandre; DIEHL, Leandro Arthur; PRADO, Fabrizio Almeida. **O que é essa tal de sensibilidade e especificidade?** 2020. Raciocínio Clínico. Disponível em: <<https://raciocinioclinico.com.br/o-que-e-essa-tal-de-sensibilidade-e-especificidade-2/>>. Acesso em: 17 jul. 2022.

BOWKER, Lindsay Newland; CHAMBERS, David M.. **The Risk, Public Liability & Economics of Tailings Storage Facility Failure**. [Washington, Eua]: Earthworks, 2015. 56 p. Disponível em: <https://earthworks.org/releases/fatal_brazilian_mine_waste_disaster_shows_modern_mining_is_increasingly_dan/>. Acesso em: 26 jul. 2022.

BRASIL. Lei nº 8.742, de 7 de dezembro de 1993. Dispõe sobre a organização da Assistência Social e dá outras providências. Lei Orgânica da Assistência Social. Brasília, Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/L8742.htm>. Acesso em: 9 jul. 2022.

BRASIL. Ministério De Estado Do Desenvolvimento Social E Combate À Fome. Portaria nº 7, de 30 de janeiro de 2012. Dispõe sobre o apoio financeiro à gestão descentralizada dos serviços, programas, projetos e benefícios de assistência social, por meio do Índice de Gestão Descentralizada do Sistema Único de Assistência Social – IGDSUAS, e dá outras providências.. Brasília, DF, Alterada pela Portaria nº 669, de 10 de setembro de 2021. Disponível em: <<http://blog.mds.gov.br/redesuas/portaria-no-7-de-30-de-janeiro-de-2012/>>. Acesso em: 20 jul. 2022.

BRASIL. Caixa Econômica Federal. Ministério da Economia. **Cadastro Único**: uma iniciativa do governo federal para identificar e conhecer as famílias brasileiras de baixa renda. Uma iniciativa do Governo Federal para identificar e conhecer as famílias brasileiras de baixa renda. [2016]. Disponível em: <<https://www.caixa.gov.br/servicos/cadastro-unico/Paginas/default.aspx>>. Acesso em: 19 jul. 2022.

BRASIL. Tribunal Regional Federal da Primeira Região. Termo de Transação e de Ajustamento de Conduta. União, Estado de Minas Gerais e Estado do Espírito Santo. Samarco SA, Vale SA, BHP Billiton Brasil LTDA. Brasília, DF, 02 de março de 2016. Termo de Transação e de Ajustamento de Conduta. Brasília, . Disponível em: <<https://www.fundacaorenova.org/arquivos-e-relatorios>>. Acesso em: 19 jul. 2022.

BRASIL. Portal Brasileiro de Dados Abertos. Controladoria-Geral da União. Microdados amostrais do Cadastro Único. [2017]. Disponível em: <<https://dados.gov.br/dataset/microdados-amostrais-do-cadastro-unico>>. Acesso em: 19 jul. 2022.

BRASIL. Ministério Público Federal. Termo de Ajustamento de Conduta nº 1. Belo Horizonte, MG, 25 de junho de 2018. Tac Governança. Belo Horizonte, . Partes: Ministério Público; Defensoria Pública; Poder Público; Empresas: Samarco, Vale e BHP Billiton Brasil; Interveniente-anuente: Fundação Renova. Disponível em: <<http://www.mpf.mp.br/grandes-casos/caso-samarco/duvidas-sobre-o-tac-governanca>>. Acesso em: 19 jul. 2022.

BRASIL. Procuradoria Geral da República. Parecer Técnico nº 279, Avaliação Participativa da Minuta do Termo de Ajustamento de Conduta Sobre A Governança (Tac-Governança), do Processo de Reparação e Recuperação dos Danos Decorrentes do Rompimento de Barragens de Rejeitos das Mineradoras Samarco, Bhp e Vale em Mariana, Minas Gerais. Solicitantes: Força Tarefa Rio Doce – MPF, Força Tarefa MPMG – Caso Samarco. Relator: Analistas do MPU Beatriz Accioly Vaz, Jorge Bruno Souza e Marco Paulo Fróes Schettino, Analistas do MPMG Camila Mattarelli de Abreu e Silva, Jonas Vaz Leandro Leal, e Assessor do MPMG Luiz Tarcizio Gonzaga de Oliveira. Parecer Nº 279/2018/Sppea. Brasília: Portal da Transparência e Prestação de Contas, 22 mar. 2018. PGR-00150655/2018. Referência Auto nº 1.22.000.003765/2015-73. Disponível em: <<http://www.transparencia.mpf.mp.br/validacaodocumento>>. Acesso em: 19 jul. 2022.

BRASIL. Lei nº 14.284, de 29 de dezembro de 2021. Institui o Programa Auxílio Brasil e o Programa Alimenta Brasil; define metas para taxas de pobreza. Brasília, Disponível em: <https://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2021/Lei/L14284.htm>. Acesso em: 01 ago. 2022.

BRASIL. Decreto nº 11.016, de 29 de março de 2022. Regulamenta o Cadastro Único para Programas Sociais do Governo Federal, instituído pelo art. 6º-F da Lei nº 8.742, de 7 de dezembro de 1993. Brasília, Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2022/decreto/D11016.htm>. Acesso em: 19 jul. 2022

BRASIL. MINISTÉRIO PÚBLICO FEDERAL. Caso Samarco. Disponível em: <<http://www.mpf.mp.br/grandes-casos/caso-samarco>>. Acesso em: 19 jul. 2022.

BRASIL. Saps - Secretaria de Atenção Primária À Saude. Ministério da Saúde. **Estratégia Saúde da Família (ESF)**. Disponível em: <<https://aps.saude.gov.br/ape/esf/>>. Acesso em: 20 jul. 2022.

BRUNBORG, Helge; URDAL, Henrik. The Demography of Conflict and Violence: an introduction. **Journal Of Peace Research**: Special Issue on the Demography of Conflict and Violence, [S.L.], v. 42, n. 4, p. 371-374, 1 jul. 2005. SAGE Publications. Disponível em: <<http://dx.doi.org/10.1177/0022343305054084>>. Acesso em: 26 jul. 2022.

COSTA, Marco Aurélio; MARGUTI, Bárbara Oliveira (ed.). **Atlas da vulnerabilidade social nos municípios brasileiros**. Brasília: Ipea, 2015. 77 p. ISBN: 978-85-7811-255-4. Disponível em: <<http://ivs.ipea.gov.br/index.php/pt/biblioteca>>. Acesso em: 27 jul. 2022.

CHAKRAVARTY, Satya R; D'AMBROSIO, Conchita. A family of unit consistent multidimensional poverty indices. In: BRESSON, Florent; BERENGER, Valerie (ed.). **Poverty and Social Exclusion Around the Mediterranean Sea**. 2013. ed. Países Baixos: Springer Us, 2013. Cap. 3. p. 75-88. Disponível em: <[10.1007/978-1-4614-5263-8_3](https://doi.org/10.1007/978-1-4614-5263-8_3)>. Acesso em: 01 jul. 2022.

DATA CAMP. **KNN Classification Tutorial using Sklearn Python**: learn k-nearest neighbor(knn) classification and build knn classifier using python scikit-learn package. Learn K-Nearest Neighbor (KNN) Classification and build KNN classifier using Python Scikit-learn package. [2018]. Disponível em: <<https://www.datacamp.com/tutorial/k-nearest-neighbor-classification-scikit-learn>>. Acesso em: 15 jul. 2022.

DIHANSTER, Willian. **KNN**: uma introdução aos algoritmos de aprendizado de máquina. uma introdução aos algoritmos de aprendizado de máquina. 2020. Dafiti Group. Disponível em: <<https://medium.com/dftblog/knn-introdu%C3%A7%C3%A3o-aos-algoritmos-de-aprendizado-de-m%C3%A1quina-dd2107693651>>. Acesso em: 15 jul. 2022.

EUROSTAT. União Europeia. **EU statistics on income and living conditions (EU-SILC) methodology - in-work poverty**. 2014. ISSN 2443-8219. Disponível em: <[https://ec.europa.eu/eurostat/statistics-explained/index.php?title=EU_statistics_on_income_and_living_conditions_\(EU-SILC\)_methodology_-_in-work_poverty](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=EU_statistics_on_income_and_living_conditions_(EU-SILC)_methodology_-_in-work_poverty)>. Acesso em: 02 jul. 2022.

FERRERO, Carlos Andres. **Algoritmo kNN para previsão de dados temporais**: funções de previsão e critérios de seleção de vizinhos próximos aplicados a variáveis ambientais em limnologia. 2009. 105 f. Dissertação (Mestrado) - Curso de Mestrado em Ciências de Computação e Matemática Computacional, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2009. Cap. 4. Disponível em: <[doi:10.11606/D.55.2009.tde-19052009-135128](https://doi.org/10.11606/D.55.2009.tde-19052009-135128)>. Acesso em: 15 jul. 2009.

FJP – FUNDAÇÃO JOÃO PINHEIRO (Minas Gerais). Seplag - Secretaria de Estado de Planejamento e Gestão (comp.). Sobre o IMRS. [2021]. Disponível em: <<http://imrs.fjp.mg.gov.br/Home/IMRS>>. Acesso em: 19 jul. 2022.

FUNDAÇÃO RENOVA. **Prefeituras de 39 municípios estão aptas a receber R\$ 500 milhões para obras de saneamento**: verba será repassada pela fundação renova, por meio do bdmg e bandes, às cidades impactadas pelo rompimento da barragem de fundão. Verba será repassada pela Fundação Renova, por meio do BDMG e Bandes, às cidades impactadas pelo rompimento da barragem de Fundão. 2018. Disponível em: <<https://www.fundacaorenova.org/release/prefeituras-de-39-municipios-estao-aptas-a-receber-r-500-milhoes-para-obras-de-saneamento/>>. Acesso em: 19 jul. 2022.

FUNDAÇÃO RENOVA. Estatuto nº 1, de 2019. Estatuto da Fundação Renova. Belo Horizonte, MG, 10 jun. 2019. Disponível em: <<https://www.fundacaorenova.org/wp-content/uploads/2020/07/estatuto-registrado-2019.pdf>>. Acesso em: 21 jul. 2022.

FUNDAÇÃO RENOVA. **Fundação Renova**: reparar o impacto do rompimento da barragem de fundão em mariana, mg. Reparar o impacto do rompimento da barragem de Fundão em Mariana, MG. Disponível em: <<https://www.fundacaorenova.org/>>. Acesso em: 20 jul. 2022.

FURTADO, Bernardo Alves. **Índice de vulnerabilidade das famílias**: atualização (2003-2009) e recortes geográficos. 1699. ed. Brasília: Ipea, 2012. 34 p. (Texto para discussão). Disponível em: <<https://www.researchgate.net/publication/235934625>>. Acesso em: 02 jul. 2022.

FURTADO, Bernardo Alves. Índice de Vulnerabilidade das Famílias (2000 - 2010): resultados. 1835. ed. Brasília: Ipea, 2013. 44 p. (Texto para discussão). Disponível em: <https://www.ipea.gov.br/portal/index.php?option=com_content&view=article&id=18165:td-1835-indice-de-vulnerabilidade-das-familias-2000-2010-resultados>. Acesso em: 02 jul. 2022.

GAMMARANO, Rosina. **The Working Poor or How a Job is no Guarantee of Decent Living Conditions**: a study based on ILO's global estimates of employment by economic class. 6. ed. [S. I.]: Ilostat, 2019. (Spotlight on work statistics). Disponível em: <<https://ilostat.ilo.org/the-working-poor-or-how-a-job-is-no-guarantee-of-decent-living-conditions/>>. Acesso em: 09 jul. 2022.

GERTLER, Paul J. *et al.* **Avaliação de impacto na prática**. 2. ed. Washington, Dc: Banco Interamericano de Desenvolvimento e Banco Mundial, 2018. 406 p. Licença: Creative Commons Attribution CC BY 3.0 IGO. Disponível em: <<https://publications.iadb.org/pt/avaliacao-de-impacto-na-pratica-segunda-edicao>>. Acesso em: 04 jul. 2022.

GOMES, Cavaliéri; GUEDES, Gilvan Ramalho. Mobilidade e vulnerabilidade socioambiental: um estudo de caso para governador valadares, minas gerais, brasil. Revista Espinhaço, [S.L.], v. 2, n. 9, p. 1-22, 11 jan. 2021. Semestral. Julho-Dezembro, 2020. ISSN: 2317-0611. Disponível em: <<http://dx.doi.org/10.5281/ZENODO.4432776>>. Acesso em: 26 jul. 2022.

GURUCHARAN, M K. **Machine Learning Basics: K-Nearest Neighbors Classification**: learn knn classification and build a model on a simple dataset to visualize your results. Learn KNN Classification and build a model on a simple dataset to visualize your results. 2020. Towards data science. Disponível em: <<https://towardsdatascience.com/machine-learning-basics-k-nearest-neighbors-classification-6c1e0b2095>>. Acesso em: 15 jul. 2022.

HORWOOD, Christopher; REITANO, Tuesday. **A Perfect Storm?**: forces shaping modern migration & displacement. [Copenhagen]: Dro, 2016. 20 p. (RMMS Discussion paper nº 3). This research is a collaboration between the Global Initiative and RMMS - the Regional Mixed Migration Secretariat. Disponível em: <<https://globalinitiative.net/analysis/a-perfect-storm-forces-shaping-modern-migration-displacement/>>. Acesso em: 25 jul. 2018.

ILOSTAT. OIT – Organização Internacional do Trabalho. Ilostat data explorer. 2019. Disponível em: <<https://www.ilo.org/shinyapps/bulkexplorer7/>>. Acesso em: 02 jul. 2022.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Ministério da Economia. **Censo Demográfico**. 2022. Portal dedicado a todos os censos, que será atualizado em breve, dado o início recente do Censo 2022. Disponível em: <<https://www.ibge.gov.br/estatisticas/sociais/populacao/22827-censo-2020-censo4.html>>. Acesso em: 26 jul. 2022.

IBM CLOUD EDUCATION. Ibm - International Business Machines Corporation. **What is Supervised Learning?** [2020]. Disponível em: <<https://www.ibm.com/cloud/learn/supervised-learning>>. Acesso em: 15 jul. 2022.

IPEA – INSTITUTO DE PESQUISA ECONÔMICA APLICADA. Ministério da Economia (ed.). Descrição das variáveis da amostra do Censo. 2010. Disponível em: <http://www.ipea.gov.br/redeipea/images/pdfs/descricao_das_variaveis_censo_2010.pdf>. Acesso em: 28 jul. 2022.

IPEA - INSTITUTO DE PESQUISA ECONÔMICA APLICADA. Ministério da Economia. ODS 1: erradicação da pobreza. Erradicação da Pobreza. 2019. Disponível em: <<https://www.ipea.gov.br/ods/ods1.html>>. Acesso em: 02 jul. 2022.

J-PAL - ABDUL LATIF JAMEEL POVERTY ACTION LAB (Cambridge) (org.). **Policy insights**. 2019. Abdul Latif Jameel Poverty Action Lab. Disponível em: <<https://www.povertyactionlab.org/policy-insights>>. Acesso em: 02 jul. 2022.

KARÁCSONYI, Dávid; TAYLOR, Andrew; BIRD, Deanne (ed.). **The Demography of Disasters: impacts for population and place**. [Nova Iorque, Eua]: Springer, 2020. 268 p. Disponível em: <<https://doi.org/10.1007/978-3-030-49920-4>>. Acesso em: 26 jul. 2022.

KARÁCSONYI, Dávid; TAYLOR, Andrew. Introduction: conceptualising the demography of disasters. In: KARÁCSONYI, Dávid; TAYLOR, Andrew; BIRD, Deanne (ed.). **The Demography of Disasters: impacts for population and place**. [Nova Iorque, Eua]: Springer, 2020. Cap. 1. p. 1-13. Disponível em: <https://doi.org/10.1007/978-3-030-49920-4_9>. Acesso em: 26 jul. 2022.

LIVI-BACCI, Massimo. Crises de mortalidade: definição, métodos de cálculo, análise das consequências. In: MARCILIO, Maria Luiza (org.). **População e sociedade: evolução das sociedades pré-industriais**. Petrópolis: Vozes, 1984. p. 69-109.

MEDEIROS, Juliana. **O que é o IGD-SUAS e qual a sua importância?** 2020. Disponível em: <<https://www.gesuas.com.br/blog/igd-suas/>>. Acesso em: 20 jul. 2022.

MINAS GERAIS. SEDRU - SECRETARIA DE ESTADO DE DESENVOLVIMENTO REGIONAL, POLÍTICA URBANA E GESTÃO METROPOLITANA. . **Avaliação dos efeitos e desdobramentos do rompimento da Barragem de Fundão em Mariana-MG**. Belo Horizonte: Sedru, 2016. 287 p. Responsável: Grupo da Força-Tarefa. Decreto nº 46.892/2015.

MONTEIRO, Simone Rocha da Rocha Pires. O marco conceitual da vulnerabilidade social. **Sociedade em Debate**, [S.I.], v. 2, n. 17, p. 29-40, 10 dez. 2010. Disponível em: <<https://revistas.ucpel.edu.br/rsd/article/view/695>>. Acesso em: 01 jul. 2022.

NOBEL PRIZE OUTREACH AB. The Nobel Foudation. **The Prize in Economic Sciences 2019**. 2019. Disponível em: <<https://www.nobelprize.org/prizes/economic-sciences/2019/summary/>>. Acesso em: 18 jul. 2022.

OCDE - ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. **Working hours needed to exit poverty**. 2022. Doi: 10.1787/77505058-en. Disponível em: <<https://data.oecd.org/benwage/working-hours-needed-to-exit-poverty.htm>>. Acesso em: 10 jul. 2022.

OIT - ORGANIZAÇÃO INTERNACIONAL DO TRABALHO. **Decent work and the sustainable development goals**: a guidebook on sdg labour market indicators. Genebra, Suíça: Oit, 2018. 92 p. Disponível em:

<https://www.ilo.org/wcmsp5/groups/public/%E2%80%94dgreports/%E2%80%94stat/documents/publication/wcms_647109.pdf>. Acesso em: 22 jul. 2022.

ONU - ORGANIZAÇÃO DAS NAÇÕES UNIDAS. **Objetivo de Desenvolvimento Sustentável**: erradicação da pobreza. Erradicação da pobreza. [2020]. Disponível em: <<https://brasil.un.org/pt-br/sdgs/1>>. Acesso em: 07 jul. 2022.

OPHI - OXFORD POVERTY AND HUMAN DEVELOPMENT INITIATIVE. University Of Oxford. **Alkire-Foster Method**. 2015. Disponível em: <<https://ophi.org.uk/research/multidimensional-poverty/alkire-foster-method/>>. Acesso em: 28 jul. 2022.

PEERSMAN, Greet. **Impact evaluation**: betterevaluation. BetterEvaluation. 2015. The content is based on ‘UNICEF Methodological Briefs for Impact Evaluation’, a collaborative project between the UNICEF Office of Research – Innocenti, BetterEvaluation, RMIT University and the International Initiative for Impact Evaluation (3ie). The briefs were written by (in alphabetical order): E. Jane Davidson, Thomas de Hoop, Delwyn Goodrick, Irene Guijt, Bronwen McDonald, Greet Peersman, Patricia Rogers, Shagun Sabarwal, Howard White. Disponível em: <<http://www.betterevaluation.org/themes/i>>. Acesso em: 27 jul. 2022.

PINKER, Steven. **Enlightenment now**: the case for reason, science, humanism and progress. Nova Iorque: Penguin Random House Llc, 2018. 576 p.

PNUD - PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO. ONU - Organização das Nações Unidas. **Human development report 2006**: beyond scarcity power, poverty and the global water crisis. Nova Iorque, Eua: Pnud, 2006. 440 p. (Human development report). Disponível em: <<https://hdr.undp.org/content/human-development-report-2006>>. Acesso em: 27 jul. 2022.

PNUD - PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO. Onu - Organização das Nações Unidas. **The Human Poverty Index (HPI)**. [2012?]. Esta página foi registrada via Wayback Machine, com este exato texto acessado em 15 de jan. de 2012..

Disponível em:

<<https://web.archive.org/web/20120115005625/http://hdr.undp.org/en/statistics/understanding/indices/hpi/>>. Acesso em: 27 jul. 2022.

PNUD - PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO. ONU - Organização das Nações Unidas. **Development approaches to displacement**. Nova Iorque: Pnud, 2016. 20 p. (Discussion papers). May, 2016. Disponível em: <<https://www.undp.org/publications/discussion-paper-development-approaches-migration>>. Acesso em: 25 jul. 2018.

PNUD – PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO. ONU – Organização das Nações Unidas. **Global Multidimensional Poverty Index 2019**: illuminating inequalities. [Nova Iorque]: Pnud, 2019. 21 p. Disponível em: <<https://hdr.undp.org/content/2019-global-multidimensional-poverty-index-mpi>>. Acesso em: 02 jul. 2022.

POLIZZI, Antonino; STRUFFOLINO, Emanuela; VAN WINKLE, Zachary. Family demographic processes and in-work poverty: a systematic review. **Advances In Life Course Research**, [S.L.], v. 52, n. 100462, p. 1-65, jun. 2022. Elsevier BV. Disponível em: <<http://dx.doi.org/10.1016/j.alcr.2022.100462>>. Acesso em: 10 jul. 2022

RASCHKA, Sebastian. **Machine Learning FAQ: why is nearest neighbor a lazy algorithm?**. Why is Nearest Neighbor a Lazy Algorithm? [2016]. Disponível em: <<https://sebastianraschka.com/faq/docs/lazy-knn.html>>. Acesso em: 15 jul. 2022.

REZENDE, Débora Borges Alves. **Avaliação de impacto do rompimento da barragem de Fundão sobre a migração dos municípios atingidos**. 2021. 122 f. Dissertação (Mestrado) - Curso de Demografia, Cedeplar - Centro de Desenvolvimento e Planejamento Regional da Faculdade de Ciências Econômicas, Universidade Federal de Minas Gerais, Belo Horizonte, 2021.

SEN, Amartya. Poverty: an ordinal approach to measurement. **Econometrica: Journal of the Econometric Society**, [S.L.], v. 44, n. 2, p. 219, mar. 1976. JSTOR. Disponível em: <<http://dx.doi.org/10.2307/1912718>>. Acesso em: 07 jul. 2022.

SEN, Amartya. **Commodities and Capabilities**. 1999. ed. Nova Déli, Índia: Oxford University Press, 1987. 89 p. Disponível em: <<https://scholar.harvard.edu/sen/publications/commodities-and-capabilities>>. Acesso em: 27 jul. 2022.

STANTON, Elizabeth A. **The Human Development Index: a history**. Amherst, EUA: Department Of Economics At The University Of Massachusetts-Amherst, 2007. 37 p. (Working paper series nº 127). Disponível em: <<https://doi.org/10.7275/1282621>>. Acesso em: 27 jul. 2022.

THE ECONOMIST. The Economist Newspaper Limited. **Towards the end of poverty: nearly 1 billion people have been taken out of extreme poverty in 20 years**. nearly 1 billion people have been taken out of extreme poverty in 20 years. 2013. Revista The Economist [online], 1. jun. 2013. Disponível em: <<https://www.economist.com/leaders/2013/06/01/towards-the-end-of-poverty>>. Acesso em: 01 jul. 2022.

UNICEF - FUNDO DAS NAÇÕES UNIDAS PARA A INFÂNCIA. **A Unicef Guide for Monitoring and Evaluation: making a difference?**. [Nova Iorque]: Unicef, 1995. 92 p. Última atualização em 25 de janeiro de 2021. Disponível em: <<https://www.corecommitments.unicef.org/kp/a-unicef-guide-for-monitoring-and-evaluation-making-a-difference>>. Acesso em: 18 jul. 2022.

WORLD DATA LAB (WDL) (Austria). **World poverty clock: explore poverty data by age, gender, and region for every country in the world until 2030**. Explore poverty data by age, gender, and region for every country in the world until 2030. 2022. Disponível em: <<https://worldpoverty.io/>>. Acesso em: 15 ago. 2022.

ANEXO I

Seguem nas próximas páginas as variáveis da base de dados desidentificada do CadÚnico, utilizadas na elaboração deste trabalho, bem como uma cópia do questionário original respondido pelo representante de uma família que precisa se inscrever no cadastro.